

# **Interspecies interaction and diversity of green sulfur bacteria**

Dissertation der Fakultät für Biologie  
der Ludwig-Maximilians-Universität München



vorgelegt von

**Johannes Friedrich Müller**

am 10 Juli 2012

1. Gutachter: Prof. Dr. Jörg Overmann, DSMZ Braunschweig

2. Gutachter: Prof. Dr. Anton Hartmann, LMU München

Tag des Promotionskolloquiums: 11.12.2012

„Wer sich Steine zurechtlegen kann, über die er stolpert, hat Erfolg in den Naturwissenschaften.“

Erwin Chargaff, österreichisch.-amerikanischer Biochemiker u. Schriftsteller (1905 - 2002)



## **Publications originating from this thesis**

### **Publication 1:**

**Müller, J.F.**, Wenter, R., Manske, A.K., Sikorski, J., Teeling, H., Garcia-Gil, L.J., Overmann, J. (2012) Biodiversity and phylogeny of the family *Chlorobiaceae* based on analyses of different genomic regions. (manuscript)

### **Publication 2:**

**Müller, J.F.**, Overmann, J. (2011) Close interspecies interactions between prokaryotes from sulfurous environments. *Front Microbiol* 2:146. (review)

### **Publication 3:**

**Müller, J.F.**, Eisenreich, W., Bunk, B., Liu, Z., Henke, P., McGlynn, S., Wenter, R., Orphan, J.V., Bryant, D.A., Overmann, J. (2012) The complex exchange of metabolites in the phototrophic consortium "*Chlorochromatium aggregatum*". (manuscript)

## Contributions of Johannes Müller to this thesis

Johannes Müller cultivated the consortia enrichment and the *Chlorobium chlorochromatii* CaD3<sup>T</sup> cultures used. He established and performed the magnetic capture experiments. Johannes Müller prepared the samples for isotopologue analysis by incubating the cultures and separating the consortia by cesium density gradient centrifugation. The subsequent analyses were conducted by Dr. Wolfgang Eisenreich. Johannes Müller prepared the central bacterium DNA used for genome analysis carried out by Prof. Dr. Donald A. Bryant and Dr. Zhenfeng Liu by separating the central bacterium from the epibiont by cesium density gradient centrifugation and extracting the DNA. He prepared the cDNA utilized for transcriptome analysis. Bioinformatic analysis of the results was performed together with Boyke Bunk who also processed the transcriptome raw data. Johannes Müller performed the fluorescence recovery after photobleaching (FRAP) analysis with the aid of Andreas Binder. Petra Henke performed the NanoSIMS analysis, Roland Wenter the amino acid excretion profile of the epibiont. Johannes Müller and Jörg Overmann wrote the review.

Johannes Müller calculated and constructed the 16S rRNA gene tree of green sulfur bacteria. The phylogenetic trees for the ITS region, the *bchg*, *sigA* and *fmoA* gene sequences of green sulfur bacteria were constructed using the pre-alignments of Ann Manske and Roland Wenter as base material. Johannes Müller also calculated and constructed the concatenated phylogenetic tree. He prepared the sequence alignments for the distance matrix comparison and the pair-wise sequence dissimilarity analysis and calculation of *ka* and *ks* values. The respective calculations and figures were supplied by Johannes Sikorski. Johannes Müller performed the coverage and diversity estimates and constructed the world map containing the sampling sites of the sequences used in the 16S rDNA tree.

I herby confirm the above statements

---

Johannes Müller

---

Prof. Dr. Jörg Overmann



# Contents

<b>Chapter 1: Summary</b>	<b>1</b>
<b>Chapter 2: Introduction</b>	<b>5</b>
2.1 Green sulfur bacteria	5
2.2 Phylogeny of green sulfur bacteria	6
2.3 Niche formation by biotic interaction	9
2.4 Characterization of phototrophic consortia	11
2.5 " <i>Chlorochromatium aggregatum</i> " as a model system for symbiotic niche formation	14
2.6 Aims of this study	17
<b>Chapter 3: Experimental procedures</b>	<b>24</b>
3.1 Experimental procedures for phylogenetic analysis of green sulfur bacteria	24
<i>Cultures of green sulfur bacteria</i>	24
<i>DNA-extraction from cultures and water samples</i>	24
<i>Standard conditions for PCR</i>	24
<i>Amplification of 16S rRNA gene sequences</i>	25
<i>PCR and cloning of the 16S-23S rRNA intergenic spacer (ITS) region</i>	27
<i>Primer design and PCR for group 1 <math>\sigma^{70}</math>-type sigma factor gene sequences (sigA)</i>	27
<i>BchG gene sequences</i>	28
<i>FmoA gene sequences</i>	28
<i>Sequencing of PCR products</i>	28
<i>Phylogenetic trees</i>	28
<i>Concatenated phylogenetic tree</i>	29
<i>Distance matrix comparisons (Mantel test)</i>	29
<i>Coverage and diversity estimates</i>	30
<i>Pair-wise sequence dissimilarity analysis and calculation of <math>k_a</math> and <math>k_s</math> values</i>	30
<i>Map of sampling sites</i>	30
<i>Nucleotide accession numbers</i>	31
3.2 Experimental procedures to investigate the molecular basis of the interspecies interaction in " <i>C. aggregatum</i> "	31
<i>Bacterial cultures and growth conditions</i>	31
<i>Probes</i>	31



<i>Nanoscale secondary ion mass spectrometry (NanoSIMS)</i> .....	32
<i>Magnetic bead separation of <sup>14</sup>C - labeled RNA</i> .....	33
<i>Cell counting</i> .....	35
<i>Fluorescence-Recovery-After-Photobleaching (FRAP)-analysis</i> .....	36
<i>Sample preparation for Isotopologomics</i> .....	36
<i>Hydrolysis of cellular protein and silylation of the resulted amino acids</i> .....	37
<i>Gas chromatography - mass spectrometry (GC/MS)</i> .....	37
<i>Genome analysis</i> .....	39
<i>In silico subtractive hybridization</i> .....	39
<i>Transcriptome analysis</i> .....	40
<i>HPLC analysis of amino acid excretion</i> .....	41
<b>Chapter 4: Results</b> .....	<b>47</b>
4.1 Results of the phylogenetic analysis.....	47
<i>16S rRNA gene sequence tree of GSB</i> .....	47
<i>Comparison of the concatenated tree with the 16S rDNA gene tree</i> .....	48
<i>Distance matrix comparisons (Mantel test)</i> .....	51
<i>Coverage and diversity estimates</i> .....	51
<i>Pair-wise sequence dissimilarity analysis and calculation of <i>k<sub>a</sub></i> and <i>k<sub>s</sub></i> values</i> .....	52
<i>World map of sampling sites</i> .....	54
4.2 Results for the investigation of the molecular basis of the interspecies interaction in “ <i>C. aggregatum</i> ” .....	56
<i>Transfer of <sup>14</sup>C with magnetic bead capture</i> .....	56
<i>NanoSIMS analysis</i> .....	60
<i>Isotopologue analysis</i> .....	60
<i>Excretion of amino acids by epibiont cells</i> .....	64
<i>Fluorescence-Recovery-After-Photobleaching (FRAP) analysis</i> .....	65
<i>Genome analysis</i> .....	66
<i>In silico subtractive hybridization</i> .....	76
<i>Transcriptome analysis</i> .....	77
<b>Chapter 5: Discussion</b> .....	<b>88</b>
5.1 Phylogeny of the green sulfur bacteria.....	88
5.1.1 Diversification of green sulfur bacteria.....	90
5.1.2 Analysis of the marker genes and the ITS region.....	93

5.1.3 Future sequencing of green sulfur bacterial strains.....	95
5.2 Molecular basis of the symbiosis in the phototrophic consortium “ <i>C. aggregatum</i> ” .....	96
5.2.1 Transfer of metabolites from the epibiont to the central bacterium.....	96
5.2.2 Mechanisms of metabolite exchange.....	101
5.2.3 Dependence of the central bacterium on the epibiont.....	103
5.2.4 Evidence of metabolite transfer from the central bacterium to the epibiont.....	105
5.2.5 Implications for the origin of phototrophic consortia.....	105
5.2.6 Features of the epibiont that relate to symbiosis.....	107
5.3 Conclusions.....	111
<b>I. Danksagung .....</b>	<b>123</b>
<b>II. Curriculum vitae .....</b>	<b>125</b>
<b>III. Supplementary material.....</b>	<b>127</b>



# Chapter 1

## Summary

The following work is shedding light on the phylogenetic classification on the family of the *Chlorobiaceae*, the members of which are showing signs of preadaptation to symbiosis. Symbioses consisting of purely prokaryotic associations between phylogenetically distinct bacterial species have been widely documented. Only few are available as a laboratory culture to elucidate the molecular basis of their interaction. One of these few model organisms is the phototrophic consortium “*Chlorochromatium aggregatum*”. It consists of 12-20 green sulfur bacteria epibionts surrounding a central, *Betaproteobacterium* in a highly ordered fashion. The phototrophic partner bacterium, belonging to the green sulfur bacteria, is available in pure culture and its physiology has been studied in detail. In this work, novel insights into the physiology of the central bacterium that was previously uncharacterized are provided.

The family of the *Chlorobiaceae* represents a phylogenetically coherent and isolated group within the domain Bacteria. Green sulfur bacteria are obligate photolithoautotrophs that require highly reducing conditions for growth and can utilize only a very limited number of carbon substrates. These bacteria thus inhabit a very narrow ecologic niche. For the phylogenetic studies on green sulfur bacteria, 323 16S rRNA gene sequences, including cultured species as well as environmental sequences were analysed. By rarefaction analysis and statistical projection, it was shown that the data represent nearly the whole spectrum of green sulfur bacterial species that can be found in the sampled habitats. Sequences of cultured species, however, did not even cover half of the biodiversity. In the 16S rDNA gene tree, different clusters were found that in most cases correlated with physiological adaptations of the included species. By combining all sampling sites of green sulfur bacteria in a world map, large, unsampled areas were revealed and it could be shown that in some regions, a non-random distribution of GSB occurred. The wide dispersal of green sulfur bacterial species can

be seen in sequences that were found ubiquitously all over the world. To imitate the phylogenetical relationships of whole genome analyses, a concatenated tree was constructed including 32 species and 3 different genetic regions, the *bchG* gene, the *sigA* gene and the *fmoA* gene. Comparison with the 16S rRNA gene tree showed more genetic differences between species, and led to a higher resolution and a more dependable phylogeny. A distance matrix comparison showed that the *fmoA* gene sequence has the highest correlation to the 16S rDNA of the sequences investigated. Additionally, a dissimilarity matrix revealed that the *fmoA* gene sequence provides the highest phylogenetic resolution among the sequences investigated. Therefore, we showed that the *fmoA* gene sequence is the most suitable among the sequences investigated to support the 16S rDNA phylogeny of green sulfur bacteria.

To overcome the limitation of immobility, some green sulfur bacteria have entered into a symbiosis with motile *Betaproteobacteria* in a type of multicellular association termed phototrophic consortia. Recent genomic, transcriptomic, and proteomic studies of "*C. aggregatum*" and its epibiont provided insights into the molecular basis and the origin of the stable association between the two very distantly related bacteria. However, to date the possibility of a metabolic coupling between the bacterial partners has not been investigated. The symbiotic exchange of metabolites between the two species was therefore investigated by tracking the flux of isotope-labeled CO<sub>2</sub> through the two partner organisms using NanoSIMS analysis and magnetic capture, revealing a fast and simultaneous incorporation of labeled carbon into both organisms. The transferred metabolites were identified by isotopologue profiling for which the partner cells were separated by cesium chloride density gradient centrifugation, a method which identified amino acids as one group of substrates to be transferred between the two partners. The addition of external carbon substrates inhibited the transfer between the two partners, suggesting that transporters are the means by which substrates are exchanged.

Genome sequencing revealed the central bacterium to be an aerobic or microaerophilic chemoheterotrophic bacterium. The existence of 32 PAS domains which are responsible for sensing various signals indicate that the central bacterium is responsible for the chemo- and phototactic responses of the consortium. The central bacterium possesses all traits of an autonomous organism. However, transcriptome analysis revealed the central bacterium to be inactive in the dark although external carbon sources were present. Thereby, a yet

unexplained dependence on the epibiont is revealed which indicates a complex metabolic coupling between the two symbiotic partner organisms.



## Chapter 2

### Introduction

#### 2.1 Green sulfur bacteria

Green sulfur bacteria are strictly anaerobic photolithoautotrophs. They assimilate CO<sub>2</sub> autotrophically through the reductive tricarboxylic acid cycle. Some organic carbon sources, such as acetate, propionate and pyruvate, can be assimilated during the phototrophic growth of the green sulfur bacteria in the presence of CO<sub>2</sub> or HCO<sub>3</sub><sup>-</sup>. All strains are capable of using sulfide as electron donor for anoxygenic photosynthesis. Sulfide is first oxidized to sulfur (zero valence sulfur) which is deposited extracellularly in sulfur globules. After sulfide depletion the sulfur is further oxidized to sulfate. Only a few strains are able to use thiosulfate or tetrathionate as electron donors. Green sulfur bacteria are typically immotile, the sole exception being the filament forming *Chloroherpeton thalassium* which is motile by gliding. Some species are able to control their vertical position with the formation of gas vesicles that enable them to position at a water depth supporting growth (Overmann 2001). Green sulfur bacteria possess a Gram-negative cell wall and are of spherical, ovoid, straight or vibrioid shape.

The ecological niche of green sulfur bacteria is determined by their requirements for light and reduced sulfur sources. Only in the narrow overlap of these opposing gradients growth is possible. This zone is mostly identical with the so-called chemocline, which commonly occurs where local conditions favor the formation of anoxic bottom water, usually in thermally stratified or meromictic lakes. There, blossoms of green sulfur bacteria can even dominate the bacterial community. A complete compilation of the various habitats of green sulfur bacteria can be found in Gernerden and Mas (1995).

A direct competitor in this ecological niche are the *Chromatiaceae*, that are also photolithoautotrophic and using reduced sulfur compounds as electron donor for anoxygenic photosynthesis. Defining their niche is the ability of *Chlorobiaceae* to form chlorosomes. These light-harvesting organelles measure 70 to 180 nm in length and 36 to 60 nm in width. They harbor bacteriochlorophylls, carotenoids and quinones, which are uniquely arranged (Olson 1998), with tight aggregates of either bacteriochlorophyll *c*, *d* or *e* being connected by

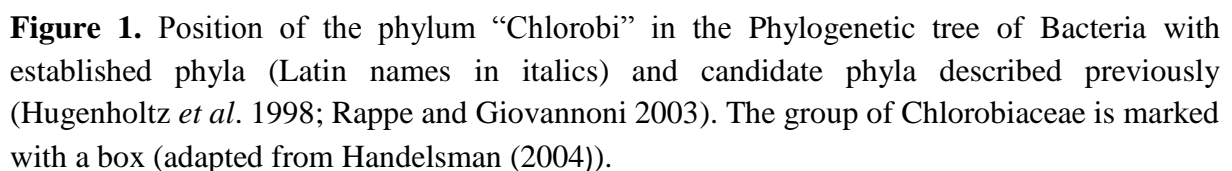


paracrystalline pigment - pigment interactions (Blankenship *et al.* 1995; Griebenow and Holzwarth 1989). In contradiction to other photosynthetic antenna complexes, the bacteriochlorophyll in chlorosomes is not attached to a protein scaffold. Thereby, up to 200.000 bacteriochlorophyll molecules can be accumulated in one chlorosome (measured in *Chlorobaculum tepidum*) (Montano *et al.* 2003). Depending on the type of bacteriochlorophyll, the green sulfur bacteria can be classified into green strains containing bacteriochlorophyll *c* or *d* and the carotenoid chlorobactene, while brown colored strains use bacteriochlorophyll *e* and isorenieratene and  $\beta$ -isorenieratene as their antenna pigments. A 5-6 nm thick base plate consisting of bacteriochlorophyll *a* and trimeric Fenna-Matthew-Olsen (FMO) proteins is coupling the chlorosomes to an iron-sulfur reaction center located in the cytoplasmic membrane (Blankenship, Madigan, and Bauer 1995). It is assumed that the role of the FMO protein is to pass citation energy from the chlorosomes to the reaction center. Per chlorosome, five to ten reaction centers are anchored in the cytoplasmic membrane (Amesz 1991). Because of the high amounts of antenna pigments compared to the small number of reaction centers, chlorosomes are highly sensitive light receptors capturing minute amounts of photons.

## 2.2 Phylogeny of green sulfur bacteria

Phylogenetically, the green sulfur bacteria form a distinct branch of the Bacteria (Garrity and Holt 2001; Overmann 2000), with non-photosynthetic Bacteroidetes as their closest relatives (Madigan *et al.* 2000). The green sulfur bacteria identified to date are very closely related bacteria. The similarity of the 16S rRNA gene of all species except the gliding filamentous marine bacterium *Chloroherpeton thalassium* that represents a phylogenetically more deeply branching member of the green sulfur bacteria (85.5 to 87 % similarity) (Overmann 2001), is  $> 90.1 \%$  ( $K_{\text{nuc}} < 0.11$ ) (Overmann 2000). The recently discovered *Ignavibacterium album* gen. nov. sp. nov. forms a new class of chemotrophic organisms in the phylum Chlorobi that lack chlorosomes and are hence unable to perform photosynthesis (Iino *et al.* 2010). The term “green sulfur bacteria” will therefore only be used for the phototrophic class Chlorobia.

The latest phylogeny of green sulfur bacteria is based only on genetic information of cultivated species (Imhoff 2003). A green sulfur bacterial phylogeny based on 16S rRNA gene sequences was first published by Figueras *et al.* (1997) and included 18 species, at that time all known species of green sulfur bacteria. They found that *Chlorobium* species can be subdivided in brown- and green-colored species and that there is a separate



7

must be the result of a difference in the functional genes. Likewise, salt tolerance is not limited to species of the marine group and closely related strains can have dissimilar salt requirements (Triado-Margarit *et al.* 2010).

Green sulfur bacteria show typical adaptations to the habitat they live in. Brown colored members of the green sulfur bacteria dominate in extreme low-light environments (Montesinos *et al.* 1983; Overmann *et al.* 1992; Overmann *et al.* 1998) and differ in the composition of their bacteriochlorophylls (Glaeser *et al.* 2002). One example is the extremely low-light adapted strain BS1 that was found in the Black Sea chemocline (Manske *et al.* 2005; Overmann *et al.* 1992), which occurs in a depth of 68-98 m. That strain is capable of photosynthetic activity at light intensities as low as  $0.015 \mu\text{mol quanta m}^{-2} \text{ s}^{-1}$  (Manske *et al.* 2005).

Green sulfur bacteria are a family of genetically closely related species. Therefore, functional genes have been used to support the 16S rDNA phylogeny and improve its resolution. A study on the importance of functional genes in the phylogeny of green sulfur bacteria was focused on the phylogenetical evaluation of *bchG* gene sequences of 14 strains of GSB (Garcia-Gil *et al.* 2003). The *bchG* gene encodes for bacteriochlorophyll *a* synthase in green sulfur bacteria and belongs to the UbiA enzyme family of polyprenyltransferases. Garcia-Gil *et al.* (2003) found that the phylogenetic relationships of green sulfur bacteria are congruent with taxonomic features, including cell shape, NaCl requirement and the presence of gas vacuoles (Overmann 2001). Evaluating the *bchG* gene sequences, they observed only very short phylogenetic distances within the GSB group, but several distinct clusters. These clusters contained brown-colored species, gas-vacuolated GSB, prosthecae GSB, or members of the genus *Chlorobium*. Like in the 16S rRNA tree (Figueras *et al.* 1997) a subdivision of saltwater strains was found (Garcia-Gil *et al.* 2003).

Group 1 sigma factors of the  $\sigma^{70}$ -type were used as a phylogenetic tool by Gruber and Bryant (Gruber and Bryant 1997) for investigation of evolutionary relationships among eubacteria. Group 1 sigma factors are the primary sigma factors in all known eubacteria and are essential for cell viability (Lonetto *et al.* 1992). The  $\sigma^{70}$ -type sigma factor is part of the eubacterial RNA polymerase holoenzyme (Burgess *et al.* 1969) and confers the promoter specificity for transcription initiation onto the core enzyme. They are useful for phylogenetic comparisons because due to structural restrictions these sigma factors possess highly conserved amino acid sequences (Helmann and Chamberlin 1988) as well as regions that are variable even among close relatives. Gruber and Bryant (1997) analyzed the phylogenetic

relationship of 45 protein sequences of group 1  $\sigma^{70}$ -type sigma factors from a variety of eubacterial groups including one sequence of green sulfur bacteria (*Chlorobaculum tepidum*). Their analysis showed a congruence of the phylogenetic clustering of sequences with the clusters found in 16S rRNA trees and an isolated position of the green sulfur bacterial sequence (Gruber and Bryant 1997).

The 16S-23S ribosomal intergenic spacer region (ITS) was shown to be a useful marker for the rapid identification of bacteria (Jensen *et al.* 1993) due to the presence of highly conserved regions as well as of highly variable regions. The region has since been used for identification of bacteria and diversity estimates in a variety of studies (Lenz *et al.* 2010; Roy *et al.* 2010; Stadthagen-Gomez *et al.* 2008).

The classical systematic assessed the diversity of *Chlorobiaceae* by using morphological attributes. The resulting system differentiated bacteria forming web-like associations (*Pelodictyon*), species forming star-shaped aggregates (*Prosthecochloris*) and green- or brown colored species (Pfennig and Trüper 1989; Pfennig 2001a, 2001b; Schlegel 1992). To date, green sulfur bacteria are subdivided into the four genera *Chlorobium*, *Prosthecochloris*, *Chlorobaculum* and *Chloroherpeton* (Imhoff 2003; Imhoff and Thiel 2010). Underlying that phylogeny is a study on 16S rDNA gene sequences and Fenna-Matthews-Olson (FMO) protein sequences of cultured GSB species (Alexander *et al.* 2002). The study included 31 strains of GSB and found a congruence of the phylogenetic trees built from 16S rDNA gene sequences and from FMO protein sequences. The clusters they found showed the incongruence of the phylogenetic trees with the current taxonomic classification of GSB species based on physiological properties, which comprised the six genera *Ancalochloris*, *Chlorobium*, *Chloroherpeton*, *Clathrochloris*, *Pelodictyon*, and *Prosthecochloris* (Overmann and Van Gemerden 2000). Consequently, a revised taxonomy of the *Chlorobiaceae* based on 16S rDNA gene and FMO protein sequences was proposed introducing new genus and species names (Imhoff 2003). The finding that the morphological attributes used to classify GSB do not match with their 16S rDNA taxonomy, leaves the question unanswered which factors led to the diversification within the family *Chlorobiaceae*. It is therefore of importance to investigate the niche formation of GSB in order to understand their phylogeny.

### 2.3 Niche formation by biotic interaction

In their natural environment, planktonic bacteria reach total cell numbers of  $10^6 \text{ ml}^{-1}$ , whereas in sediments and soils,  $10^9$  and  $10^{11}$  bacterial cells $\cdot\text{cm}^{-3}$ , respectively, have been observed

(Fægri *et al.* 1977; Whitman *et al.* 1998). Assuming a homogenous distribution, distances between bacterial cells in these environments would amount to 112  $\mu\text{m}$  for planktonic, 10  $\mu\text{m}$  for sediment environments and about 1  $\mu\text{m}$  for soil bacteria (Overmann 2001b). Taking into account the estimated number of bacterial species in soil that range from 500,000 (Dykhuizen 1998) to  $8.3 \cdot 10^6$  (Gans *et al.* 2005), the closest neighbors of each cell statistically should represent different species. A spatially close association of different bacterial species can result in metabolic complementation or other synergisms. In this context, the most extensively studied example is the conversion of cellulose to methane and carbon dioxide in anoxic habitats. The degradation is only possible by a close cooperation of at least four different groups of bacteria that encompass primary and secondary fermenting bacteria as well as two types of methanogens. Along this anaerobic food chain, end products of one group are exploited by the members downstream the flow of electrons. Although the bacteria involved in the first steps of cellulose degradation do not obligately depend on the accompanying bacteria for provision of growth substrates, they profit energetically from the rapid consumption of their excretion products. This renders their metabolism energetically more favorable or makes some reactions even possible (Bryant 1979; McInerney 1986; Schink 1992; Zehnder *et al.* 1982). Recent studies of syntrophic communities in Lake Constance profundal sediments yielded new and unexpected results. The dominant sugar-degrading bacteria were not the typical fermenting bacteria that dominate in anaerobic sludge systems or the rumen environment. They rather represented syntrophic bacteria most closely related to the genus *Bacillus* that could only be grown anaerobically and in co-culture with the hydrogen-using methanogen *Methanospirillum hungatei* (Müller *et al.* 2008). For efficient syntrophic substrate oxidation, close physical contact of the partner organisms is indispensable. Monocultures of *Pelotomaculum thermopropionicum* strain SI and *Methanothermobacter thermautotrophicus* show dispersed growth of the cells. In contrast, cocultures of the two strains formed tight aggregates when grown on propionate, for which the maximal distance for syntrophic propionate oxidation was estimated to be approximately 2  $\mu\text{m}$  (Ishii *et al.* 2005). Interestingly, the  $\text{H}_2$ -consuming partner in syntrophic relationships can be replaced by an  $\text{H}_2$ -purging culture vessel, allowing *Syntrophothermus lipocalidus* to grow on butyrate and *Aminobacterium colombiense* on alanine in pure culture (Adams *et al.* 2006). Thus, the syntrophic associations investigated to date are typically based on efficient  $\text{H}_2$ -removal as obligate basis for their interdependence. Additional types of bacterial interactions have been described more recently. Cultures of *Pseudomonas aeruginosa* were shown to only grow on chitin if in coculture with a chitin degrading bacterium like










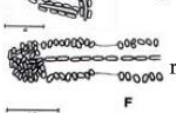
*Aeromonas hydrophila*. In addition to simply growing on the degradation products produced by the exoenzymes of the partner, *P. aeruginosa* induced release of acetate in *A. hydrophila* by inhibiting its aconitase employing pyocyanin. The resulting incomplete oxidation of chitin to acetate by *A. hydrophila* is then exploited by *P. aeruginosa* for its own growth (Jagmann et al. 2010).

Although these well characterized associations certainly are of major ecological relevance in their respective environments, they were typically obtained using standard defined growth media. As a result, presently available laboratory model systems were selected based on their ability to grow readily and - under at least some experimental conditions - on their own in pure culture. Obviously, this cultivation strategy counterselects against bacterial associations of obligately or at least tight interdependence. While the significant advances in the development of cultivation-independent techniques permit a partial analysis of so-far-uncultured associations (Blumenberg et al. 2004; Orphan et al. 2001; Pernthaler et al. 2008), laboratory grown model systems are still indispensable for in-depth studies of gene expression and metabolism.

## 2.4 Characterization of phototrophic consortia

Phototrophic consortia were already discovered in 1906 (Lauterborn 1906) and invariably encompass green or brown-colored bacteria as epibionts that surround a central bacterium in a highly ordered fashion. Several decades later, electron microscopic analyses documented the presence of chlorosomes in the epibiont cells and led to the conclusion that the phototrophic epibionts belong to the green sulfur bacteria (Caldwell and Tiedje 1975). This was confirmed by the application of fluorescence *in situ* hybridization employing a highly specific oligonucleotide probe against green sulfur bacterial 16S rRNA (Tuschak et al. 1999). The other partner bacterium of the symbiosis remained much less investigated than its epibionts. First, it had even been overlooked due to its low contrast in the light microscope. Over 90 years later, the central bacterium was identified as a *Betaproteobacterium* (Frössl and Overmann 2000) that exhibits a rod shaped morphology with tapered ends (Overmann et al. 1998). Electron microscopy revealed the cells to be monopolarly monotrichously flagellated (Glaeser and Overmann 2003b). Within the *Betaproteobacteria* the central bacterium represents a so far isolated phylogenetic lineage belonging to the family of the *Comamonadaceae*. The closest relatives are *Rhodoferrax* spp., *Polaromonas vacuolata* and *Variovorax paradoxus* (Kanzler et al. 2005).

**Table 1.** Overview of phototrophic consortia (modified from (Fröstl and Overmann 2000; Overmann and Van Gernerden 2000; Overmann 2006). n.d. = not determined

Consortium	Morphology	Colour of epibionts	Number of epibionts	Gas vesicles	References
" <i>Chlorochromatium aggregatum</i> "		green	$12.9 \pm 4.5$	-	Lauterborn 1906 Fröstl & Overmann 2000
" <i>Chlorochromatium glebulum</i> "		green	7 - 40	+	Skuja 1957 Fröstl & Overmann 2000
" <i>Chlorochromatium magnum</i> "		green	$36 \pm 4.4$	+	Fröstl & Overmann 2000
" <i>Chlorochromatium lunatum</i> "		green	$68.6 \pm 2.5$	+	Abella <i>et al.</i> 1998
" <i>Pelochromatium roseum</i> "		brown	$19.4 \pm 4.4$	-	Lauterborn 1913 Tuschak <i>et al.</i> 1999
" <i>Pelochromatium selenoides</i> "		brown	$44.5 \pm 3.5$	+	Abella <i>et al.</i> 1998
" <i>Pelochromatium roseo-viride</i> "		green/brown	n.d.	+	Gorlenko & Kusnetzov 1972
" <i>Pelochromatium latum</i> "		brown	brown	?	Glaeser & Overmann 2004
" <i>Chloroplana vacuolata</i> "		green	= 400	+	Dubinina & Kusnetzov 1976
" <i>Cylindrogloe a bacterifera</i> "		green	n.d.	-	Perfiliev 1914 Skuja 1957

Based solely on their morphology, ten different phototrophic consortia can be distinguished to date (Overmann 2001a; Overmann and Schubert 2002). The majority of the morphotypes are motile, motility being conferred by the central colorless bacterium. The 13-69 epibiont cells are either green or brown-colored representatives of the green sulfur bacteria. The smaller consortia like "*Chlorochromatium aggregatum*" (harboring green epibionts) and "*Pelochromatium roseum*" (brown epibionts), are barrel shaped and consist of 12-20 epibiont cells (Overmann *et al.* 1998). Rather globular in shape and consisting of  $\geq 40$  epibionts are the significantly larger consortia "*Chlorochromatium magnum*" (green epibionts) (Fröstl and Overmann 2000), "*Pelochromatium latum*" (brown epibionts) (Glaeser and Overmann 2004) and "*Pelochromatium roseo-viride*" (Gorlenko and Kusnezow 1972). The latter consortium is the only one harboring two types of epibionts, with brown cells forming an inner layer and green ones an outer layer. "*Chloroplana vacuolata*" and "*Cylindrogloe a bacterifera*" can be distinguished from the other consortia by their immotility and different cell arrangement.

"*Chloroplana vacuolata*" consists of rows of green sulfur bacteria alternating with colorless bacteria forming a flat sheath (Dubinina and Kuznetsov 1976), with both species containing gas vacuoles. In "*Cylindrogloea bactifera*", a slime layer containing green sulfur bacteria is surrounding filamentous, colorless bacteria (Perfiliev 1914; Skuja 1956). Since they consist of two different types of bacteria, the names of consortia are without standing in nomenclature (Trüper and Pfennig 1971) and, accordingly, are given here in quotation marks.

When 16S rRNA gene sequences of green sulfur bacteria from phototrophic consortia were investigated from a total of 14 different lakes in Europe and North America (Glaeser and Overmann 2004), a total of 19 different types of epibionts could be detected. Of those, only two types occurred on both, the European and North American continents. Although morphologically identical consortia from one lake always contained just a single epibiont phylotype, morphologically indistinguishable consortia from different lakes frequently harbored phylogenetically different epibionts. Phylogenetic analyses demonstrated that the epibiont sequences do not constitute a monophyletic group within the radiation of green sulfur bacteria. Therefore, the ability to form symbiotic interactions is a trait of GSB that allows niche formation and was gained independently by different ancestors of epibionts or, alternatively, was present in the common ancestor of the green sulfur bacteria, contributing to the diversification of the *Chlorobiaceae*.

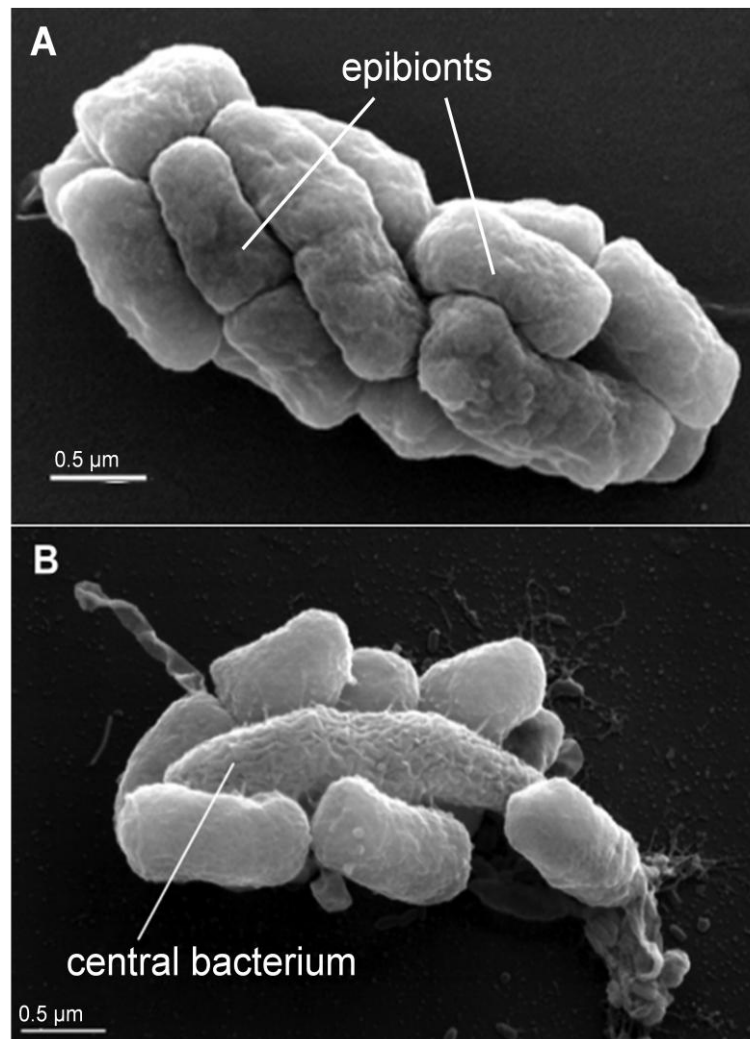
In parallel, the phylogeny of central bacteria of phototrophic consortia was investigated. This analysis exploited a rare tandem *rrn* operon arrangement in these bacteria that involves an unusual short interoperon spacer of 195 bp (Pfannes et al. 2007). *Betaproteobacteria* with this genomic feature were exclusively encountered in chemocline environments and form a novel, distinct and highly diverse subcluster within the subphylum. Within this cluster, the sequences of central bacteria of phototrophic consortia were found to be polyphyletic. Thus, like in their green sulfur bacterial counterparts, the ability to become a central bacterium may have evolved independently in several lineages of betaproteobacteria, or was already present in a common ancestor of the different central bacteria. In the barrel-shaped types of consortia, the green sulfur bacterial epibionts have overcome their immotility. However, the existence of two different types of nonmotile consortia indicates that motility is not the only advantage gained by green sulfur bacteria that form these interspecies associations with heterotrophic bacteria. As described in this work, the exchange of metabolites plays a major role in the symbiotic interaction, and might therefore be the key selective factor of symbiosis in immotile phototrophic consortia.



## 2.5 Establishing "*Chlorochromatium aggregatum*" as a model system for symbiotic niche formation

"*Chlorochromatium aggregatum*" is the only phototrophic consortium at present that can be successfully cultivated in the laboratory (Fröstl and Overmann 1998). With a stable enrichment culture, it was possible to isolate the epibiont of the consortium in pure culture by deep agar dilution series (Vogl et al 2006). After the first attempt in 1957 (Mechsner) in which the epibiont culture of "*Chlorochromatium aggregatum*" was lost before it could be studied in detail, a stable pure culture could now be established and described. On the basis of 16S rRNA sequence comparison, the strain is distantly related to other known green sulfur bacteria ( $\leq 94.6\%$  sequence homology) and therefore described as a novel species within the genus *Chlorobium*, *Chlorobium chlorochromatii* strain CaD. However, physiological and molecular analyses of the novel isolate did not provide major differences to already described strains of the phylum *Chlorobi*. *Chlorobium chlorochromatii* is obligately anaerobic and photolithoautotrophic, photoassimilating acetate and peptone in the presence of sulfide and hydrogencarbonate (Vogl, et al. 2006). Chlorosomes contain bacteriochlorophylls *a* and *c*, and  $\gamma$ -carotene and OH- $\gamma$ -carotene glucoside laurate as the dominant carotenoids. A difference to their free-living counterparts could only be described in the low cellular concentration of carotenoids in the epibiont. This slight anomaly had also been observed in the brown epibionts of the phototrophic consortium "*Pelochromatium roseum*" (Glaeser et al 2002; Glaeser and Overmann 2003a).

The other partner bacterium of the symbiosis is by far not as well investigated as its epibionts. The presence of a central bacterium had even been overlooked due to its low contrast in the light microscope when "*Chlorochromatium aggregatum*" was first described in 1906 (Lauterborn). Its morphology is rod shaped with tapered ends (Overmann, Tuschak, Sass, and Fröstl 1998) and it could be identified as monopolarly monotrichously flagellated by electron microscopy (Glaeser and Overmann 2003b). Being outnumbered around twentyfold by the epibiont and with other chemotrophic bacteria in the enrichment culture, the central bacterium eludes the established physiological and most of the molecular methods to date. Efforts to cultivate the central bacterium without its epibionts have been to no avail.



**Figure 2.** Scanning electron photomicrographs of “*Chlorochromatium aggregatum*” (courtesy of K. Vogl and Prof. Dr. G. Wanner). **A.** Intact consortium after fixation with 2 % glutardialdehyde. Epibionts are covering the central bacterium. **B.** Partially disaggregated consortium after exposure to air without fixation.

Efforts to sequence the 16S rRNA of the central bacterium have revealed that it belongs to the *Betaproteobacteria* (Pfannes 2007). Within this group, the central bacterium represents a so far isolated phylogenetic lineage belonging to the family of the *Comamonadaceae*. The closest relatives are *Rhodoferrax* spp., *Polaromonas vacuolata* and *Variovorax paradoxus*.

Due to scotophobic reaction “*Chlorochromatium aggregatum*” in enrichment culture accumulate in dim light ( $\leq 1.5 \mu\text{mol quanta}\cdot\text{m}^{-2}\cdot\text{s}^{-1}$ ) (Fröstl and Overmann 1998). Furthermore, the consortium exhibits a chemotactic behavior towards sulfide *in situ* (Fröstl and Overmann 1998; Glaeser and Overmann 2003b) and in enrichment culture (Glaeser and Overmann 2003b). Thus, the consortium can orientate itself in light and sulfide gradients,

which allows it to reach water layers of optimal conditions for photosynthesis in a short period of time. In contrast, green sulfur bacteria are immotile and only strains able to produce gas vacuoles can influence their vertical position, whereas this change in buoyant density occurs over a time period of several days (Overmann et al 1991). How phototrophic consortia use their advantage in motility over green sulfur bacteria has been observed in two meromictic lakes in Tasmania which are characterized by a sharp chemocline. In contrast to the rest of the microbial community in the chemocline, "*C. aggregatum*" has been found to vary its position from high in the microaerobic zone (0.25 m above the redoxcline), at an oxygen concentration of 0-4 mg·l<sup>-1</sup> and a redox potential of + 350 mV, to well within the anaerobic zone, with a sulfide concentration of 30 mg·l<sup>-1</sup> and a redox potential of - 100 mV (Croome and Tyler 1984; Overmann, Lehmann, and Pfenning 1991).

When the proteome of *Chlorobium chlorochromatii* CaD in the free-living state was compared to that of the symbiotic state by 2-D differential gel electrophoresis (2-D DIGE), it became apparent that symbiosis-specific regulation involves genes of central metabolic pathways rather than symbiosis-specific genes (Wenter et al. 2010). In the soluble proteome, 54 proteins were expressed exclusively in consortia. This finding is the first indication that the basis of the symbiotic interaction in consortia involves metabolic coupling of the two partner organisms. Among the 54 proteins were a considerable number of proteins involved in amino acid metabolism. These included glutamate synthase, 2-isopropylmalate synthase and the nitrogen regulatory protein P-II. The latter showed the highest overall upregulation that amounted to a 189-fold increase in transcript abundance as determined by subsequent RT-qPCR. It is thereby concluded that the amino acid requirement in the consortium is higher than in the epibiont in pure culture. Therefore, they might be candidate substrates involved in the metabolic coupling of the symbiotic partners.

Parallel investigations of the membrane proteome revealed that a branched chain amino acid ABC-transporter binding protein was expressed only in the associated state of the epibiont. Interestingly, the expression of the ABC-transporter binding protein could also be induced in the free-living state by addition of sterile filtered supernatant of the consortia culture, but not with peptone or branched chain amino acids themselves. This is an evidence for a signal exchange between the two symbiotic partners mediated through the surrounding medium.

The results of the proteome analysis were supplemented by transcriptomic studies of the epibiont in the associated and the free-living state (Wenter et al. 2010). Of the 328 differentially expressed genes, 19 genes were found to be up-regulated and are involved in

amino acid synthesis while six genes of the amino acid pathways were down-regulated. The conclusion that nitrogen metabolism of the epibiont is stimulated in the symbiotic state is commensurate with the simultaneous up-regulation of the *nifH*, *nifE*, and *nifB* genes and with the prominent expression of the P-II nitrogen regulatory protein.

## 2.6 Aims of this study

The phototrophic consortium “*C. aggregatum*” is one of the few close interspecies interactions that can be studied in the laboratory and can therefore serve as a model system for prokaryotic interactions (Overmann 2005). Furthermore, research on phototrophic consortia has shown their ecological relevance (Gasol et al. 1995). To date, genomic and proteomic (Wenter et al. 2009) analysis of the epibiont in pure culture and in association with its symbiotic partner have provided a first insight into the molecular basis of the symbiosis. However, nothing is known about the metabolic coupling within the interspecies interaction. Thus, one goal of the thesis was to investigate if and how metabolites are exchanged between the two symbiotic partners. The flow of carbon should therefore be tracked on its path through the phototrophic consortium by a series of labeling experiments. In contrast to the epibionts, the central bacteria of phototrophic consortia could neither be enriched nor isolated to date. As a consequence, a main subject of this thesis was to separate the two symbiotic partners from each other. From the separation of the two consortia species, a deeper understanding of the metabolic interactions should be gained by genome sequencing of the heterotrophic partner and separately analyzing the metabolic products of the two organisms.

Furthermore, the family of the epibiont, the *Chlorobiaceae* was to be investigated. The overall diversity of the phylum Chlorobia is yet unknown due to numerous uncultured species and various habitats of green sulfur bacteria. For analyzing the overall phylogeny of green sulfur bacteria, it is therefore indispensable to use as many genetic information as possible and to include cultured strains as well as environmental sequences.

In the phylogenetic study, multiple genetic regions as well as 16S rDNA gene sequences, from both cultured strains and environmental sequences, should therefore be employed. In this comprehensive study, it should be investigated how much of the green sulfur bacterial diversity is covered to date and how much diversity there is yet to discover. With this understanding, targets for eventual new sequencing projects should be identified and the effectiveness of the culture techniques utilized for green sulfur bacteria should be evaluated.

Since the members of the family *Chlorobiaceae* are closely related, one aim of this study was to identify a suitable marker gene with which the 16S rDNA phylogeny could be supported. Candidate gene sequences should therefore be tested for their phylogenetic resolution and for the correlation of their phylogenetic trees with the 16S rDNA tree. To further investigate the diversification of the green sulfur bacteria, a concatenated tree using three marker genes should be constructed to imitate the phylogenetical relationships of whole genome analyses.

## References

- Alexander, B., Andersen, J. H., Cox, R. P., and Imhoff, J. F. (2002). Phylogeny of green sulfur bacteria on the basis of gene sequences of 16S rRNA and of the Fenna-Matthews-Olson protein. *Arch Microbiol* 178(2), 131-40.
- Burgess, R. R., Travers, A. A., Dunn, J. J., and Bautz, E. K. (1969). Factor stimulating transcription by RNA polymerase. *Nature* 221(5175), 43-6.
- Figueras, J. B., Garcia-Gil, L. J., and Abella, C. A. (1997). Phylogeny of the genus *Chlorobium* based on 16S rDNA sequence. *FEMS Microbiology Letters* 152(1), 31-36.
- Fröstl, J. M., and Overmann, J. (2000). Phylogenetic affiliation of the bacteria that constitute phototrophic consortia. *Arch Microbiol* 174(1-2), 50-8.
- Garcia-Gil, L. J., Gich, F. B., and Fuentes-Garcia, X. (2003). A comparative study of bchG from green photosynthetic bacteria. *Archives of microbiology* 179(2), 108-115.
- Garrity, G. M., and Holt, J. G. (2001). Phylum BXI. Chlorobi phy. nov. *Bergey's Manual of Systematic Bacteriology* 1, 601-623.
- Gasol, J. M., Jurgens, K., Massana, R., Calderon-Paz, J. I., and Pedros-Alio, C. (1995). Mass developing of *Daphnia pulex* in a sulfide-rich pond (Lake Ciso). *Archiv für Hydrobiologie* 132(3), 279-296.
- Glaeser, J., Baneras, L., Rotters, H., and Overmann, J. (2002). Novel bacteriochlorophyll *a* structures and species-specific variability of pigment composition in green sulfur bacteria. *Archives of microbiology* 177(6), 475-485.
- Gruber, T. M., and Bryant, D. A. (1997). Molecular systematic studies of eubacteria, using sigma70-type sigma factors of group 1 and group 2. *J Bacteriol* 179(5), 1734-47.
- Handelsman, J. (2004). Metagenomics: Application of Genomics to Uncultured Microorganisms. *Microbiol Mol Biol Rev* 68(4), 12.
- Helmann, J. D., and Chamberlin, M. J. (1988). Structure and function of bacterial sigma factors. *Annu Rev Biochem* 57, 839-72.
- Hugenholtz, P., Pitulle, C., Hershberger, K. L., and Pace, N. R. (1998). Novel division level bacterial diversity in a Yellowstone hot spring. *Journal of bacteriology* 180(2), 366-376.

- Iino, T., Mori, K., Uchino, Y., Nakagawa, T., Harayama, S., and Suzuki, K. I. (2010). *Ignavibacterium album* gen. nov., sp. nov., a moderately thermophilic anaerobic bacterium isolated from microbial mats at a terrestrial hot spring and proposal of *Ignavibacteria* classis nov., for a novel lineage at the periphery of green sulfur bacteria. *International Journal of Systematic and Evolutionary Microbiology* 60(6), 1376-1382.
- Imhoff, J. F. (2003). Phylogenetic taxonomy of the family Chlorobiaceae on the basis of 16S rRNA and fmo (Fenna-Matthews-Olson protein) gene sequences. *International Journal of Systematic and Evolutionary Microbiology* 53(4), 941-951.
- Imhoff, J. F., and Thiel, V. (2010). Phylogeny and taxonomy of Chlorobiaceae. *Photosynth Res* 104(2-3), 123-36.
- Jensen, M. A., Webster, J. A., and Straus, N. (1993). Rapid identification of bacteria on the basis of polymerase chain reaction-amplified ribosomal DNA spacer polymorphisms. *Appl Environ Microbiol* 59(4), 945-52.
- Lenz, O., Beran, P., Fousek, J., and Mráz, I. (2010). A microarray for screening the variability of 16S–23S rRNA internal transcribed spacer in *Pseudomonas syringae*. *Journal of Microbiological Methods* 82(1), 90-94.
- Lonetto, M., Gribskov, M., and Gross, C. A. (1992). The sigma 70 family: sequence conservation and evolutionary relationships. *J Bacteriol* 174(12), 3843-9.
- Madigan, M. T., Martinko, J. M., and Parker, J., Eds. (2000). Brock – Mikrobiologie. German ed. Edited by W. Goebel. Berlin, Germany: Spektrum, Akad. Verlag
- Manske, A. K., Glaeser, J., Kuypers, M. M., and Overmann, J. (2005). Physiology and phylogeny of green sulfur bacteria forming a monospecific phototrophic assemblage at a depth of 100 meters in the Black Sea. *Appl Environ Microbiol* 71(12), 8049-60.
- Montesinos, E., Guerrero, R., Abella, C., and Esteve, I. (1983). Ecology and Physiology of the Competition for Light Between *Chlorobium limicola* and *Chlorobium phaeobacteroides* in Natural Habitats. *Appl Environ Microbiol* 46(5), 1007-16.
- Overmann, J. (2000). The family Chlorobiaceae. *The Prokaryotes. An Evolving Electronic Resource for the Microbiological Community, Vol Release 3.1*.
- Overmann, J. (2001). Green sulfur bacteria. *Bergey's Manual of Systematic Bacteriology* 1, 601-605.
- Overmann, J., Ed. (2005). Symbiosis between non-related bacteria in phototrophic consortia. Molecular basis of symbiosis. Progress in Molecular Subcellular Biology. Edited by J. Overmann. Berlin, Germany: Springer-Verlag.

- Overmann, J. (2006). Symbiosis between non-related bacteria in phototrophic consortia. *Progress in molecular and subcellular biology* 41, 21-37.
- Overmann, J., Cypionka, H., and Pfennig, N. (1992). An extremely low-light-adapted phototrophic sulfur bacterium from the Black Sea. *Limnology & Oceanography* 37(1), 150-155.
- Overmann, J., Tuschak, C., Sass, H., and Frostl, J. (1998). The ecological niche of the consortium "Pelochromatium roseum". *Arch Microbiol* 169(2), 120-8.
- Overmann, J., and Van Gernerden, H. (2000). Microbial interactions involving sulfur bacteria: Implications for the ecology and evolution of bacterial communities. *FEMS microbiology reviews* 24(5), 591-599.
- Pfannes, K. R. (2007). Characterization of the symbiotic bacterial partners in phototrophic consortia, Dissertation University of Munich. 180p
- Pfennig, N., and Trüper, H. G. (1989). Anoxygenic phototrophic bacteria. *Bergey's Manual of Systematic Bacteriology* 3, 1635-1709.
- Pfennig, N. a. O., J. (2001a). "Bergey's manual of systematic bacteriology." 2 ed. Chlorobium (D. R. Boone, R. W. Castenholz, and G. M. Garrity, Eds.), 1 Springer, Heidelberg, Germany.
- Pfennig, N. a. O., J. (2001b). "Bergey's manual of systematic bacteriology." 2 ed. Pelodictyon (D. R. Boone, R. W. Castenholz, and G. M. Garrity, Eds.), 1 Springer, Heidelberg, Germany.
- Rappe, M. S., and Giovannoni, S. J. (2003). The Uncultured Microbial Majority, Vol. 57, pp. 369-394.
- Roy, L., Dowling, A. P., Chauve, C. M., and Buronfosse, T. (2010). Diversity of phylogenetic information according to the locus and the taxonomic level: an example from a parasitic mesostigmatid mite genus. *Int J Mol Sci* 11(4), 1704-34.
- Schlegel, H. G. (1992). Allgemeine Mikrobiologie. Revised 7th edition. Thieme, Stuttgart, Germany.
- Stadthagen-Gomez, G., Helguera-Repetto, A. C., Cerna-Cortes, J. F., Goldstein, R. A., Cox, R. A., and Gonzalez-y-Merchand, J. A. (2008). The organization of two rRNA (rrn) operons of the slow-growing pathogen Mycobacterium celatum provides key insights into mycobacterial evolution. *FEMS Microbiology Letters* 280(1), 102-112.
- Triado-Margarit, X., Vila, X., and Abella, C. A. (2010). Novel green sulfur bacteria phylotypes detected in saline environments: ecophysiological characters versus phylogenetic taxonomy. *Antonie van Leeuwenhoek* 97(4), 419-31.



- Van Gemerden, H., and Mas, J. (1995). Ecology of phototrophic sulfur bacteria. *Anoxygenic Photosynthetic Bacteria*, 49-85.
- Wenter, R., Wanner, G., Schuler, D., and Overmann, J. (2009). Ultrastructure, tactic behaviour and potential for sulfate reduction of a novel multicellular magnetotactic prokaryote from North Sea sediments. *Environmental microbiology* 11(6), 1493-1505.



## Chapter 3

### Experimental procedures

#### 3.1 Experimental procedures for phylogenetic analysis of green sulfur bacteria

##### *Cultures of green sulfur bacteria*

Green sulfur bacterial strains of which sequences were obtained in this thesis (suppl. Table S1) were kept in standard saline or freshwater SL10 medium (Widdel et al. 1983). The strain BS-1 was kept in artificial seawater medium (Coolen and Overmann 2000), adjusted to the ionic strength of the Black Sea chemocline (Manske et al. 2005).

##### *DNA-extraction from cultures and water samples*

Genomic DNA of bacterial strains was isolated from exponentially growing cultures by using the DNeasy Tissue Kit (Qiagen, Hilden, Germany) following the instructions supplied by the manufacturer. DNA content was quantified using Pico Green (Molecular Probes, Eugene, USA).

##### *Standard conditions for PCR*

Standard conditions for PCR comprised 20 to 50 ng DNA or 0.5 µl of a cell suspension (cell pellet from 2 ml of culture, resuspended in 10 mM Tris·HCl, pH 8.5), 10 µM of each primer, 5 µl of GeneAMP 10x PCR buffer, 0.2 mM of each dNTP (GeneAMP dNTPs, Applied Biosystems, Weiterstadt, Germany), 3.5 mM MgCl<sub>2</sub> and 1.25 U AmpliTaq Gold polymerase (Applied Biosystems) in a total volume of 50 µl. Amplifications were performed in a GeneAMP thermo cycler PCR system 2400 or 9700 (Applied Biosystems) (Table 2). Amplification products were analyzed by standard agarose gel electrophoresis.

**Table 2.** PCR conditions and step down PCR. Each PCR started with a hot start (95°C, 5 min) and ended with a 4°C hold.

Target	Primer combination	Melting	Step down PCR: Annealing (No. of cycles)		Extension
16S rRNA gene	Uni8f – 1492r	94°, 30 s	59°C, 45 s (10)	54°C, 45 s (20)	72°C, 60 s
ITS region	GSB822f – L1r, 1525f – L1r	94°, 30 s	55°C, 60 s (10)	50°C, 60 s (25)	72°C, 60 s
Sigma factor A	GSB-SigA-F4 – GSB-SigA-R1 <sup>a</sup>	94°, 30 s	61°C, 45 s (10)	56°C, 45 s (20)	72°C, 60 s
	Sig208f – Sig827r <sup>a,b</sup>	94°, 60 s	61°C, 45 s (10)	56°C, 45 s (25)	72°C, 60 s
	Sig312f – Sig827r <sup>a,b</sup>	94°, 45 s	56°C, 45 s (10)	52°C, 45 s (25)	72°C, 60 s
BchG gene	bchG-F – bchG-R	94°, 30 s	57°C, 60 s (10)	50°C, 60 s (20)	72°C, 60 s
ERIC-PCR	ERIC-1R – ERIC-2 <sup>b</sup>	95°, 30 s	52°C, 60 s (30)	-	70°C, 8 min

<sup>a</sup> This PCR comprised an additional hold at 72°C (10 min) after the amplification cycles, before cooling to 4°C.

<sup>b</sup> This PCR initiated with a hot start (96°C, 4 min), followed by a hold (80°C, 4 min) where the polymerase was added.

### ***Amplification of 16S rRNA gene sequences***

The primer pair 8f/1492r (Table 3) was used in standard PCR conditions in a step-down PCR to amplify the almost complete 16S rRNA gene sequence (Table 2). For sequencing, the primers 8f, Uni341f, Uni515f, Uni517r, Uni 907r, 926f, 1055r, and 1492r (Table 3) were used.

**Table 3.** Primers used for phylogenetic analyses

Primer name	Sequence	Reference
<b>16S rRNA gene<sup>a</sup></b>		
8f	5'-AGA GTT TGA TCC TGG CTC AG-3'	Lane 1991
Uni341f	5'-CC TAC GGG AGG CAG CAG-3'	Muyzer et al. 1993
Uni515f	5'-GTG CCA GCA GCC GCG G-3'	Lane 1991, modified
517r	5'-ATT ACC GCG GCT GCT GGC-3'	Lane 1991
Uni907r	5'-CCG TCA ATT CCT TTG AGT TT-3'	Lane 1991
926f	5'-AAA CTC AAA GGA ATT GAC GG-3'	Lane 1991
1055r	5'-AGC TGA CGA CAG CCA T-3'	Amann et al. 1995
Uni1492r	5'-GGT TAC CTT GTT ACG ACT T-3'	Weisburg et al. 1991
<b>ITS region<sup>a</sup></b>		
1525f	5'-GG(CT) TGG A(GC)C ACC TCC TT-3'	Lane 1991
L1r	5'-CAA GGC ATC CAC CGT-3'	Jensen et al. 1993
GSB822f	5'-AAT ACT AGA TGT TGG TCA T-3'	Overmann et al. 1999
<b>Sigma factor A</b>		
Sig208f	5'-AA(CT) CT(GC) CG(CT) (CT)T(GC) GTG GT(CT) TC(GT) GTG GC-3'	Don Bryant, pers. comm.
Sig312f	5'-CAG TA(CT) CAG AA(CT) CAG GG-3'	Don Bryant, pers. comm.
Sig827r	5'-ATC TG(GC) CGG AC(CT) CG(CT) TC(GC) CG(AG) GT(AG) A-3'	Don Bryant, pers. comm.
GSB-SigA-F4	5'-ATT GTG CG(AC) (CT)T(GT) CC-3'	this study
GSB-SigA-R1	5'-AT(AT) GG(CT) ATG GA(CT) AAT CCG CT-3'	this study
<b>BChG gene</b>		
bchG-F	5'-GTC GTG ACT GGG AAA ACA TGA A(AG)C C(AGCT)G T(GCT)A C(AGTC)T GG-3'	Garcia-Gil et al. 2003
bchG-R	5'-ACG GAC TTG AAG TCG TTC ATG (AG)TC ATG A(AT)(AG) CC-3'	Garcia-Gil et al. 2003
BchG-R Seq	5'-ACG GAC TTG AAG TCG TTC ATG-3'	Frederic Gich, pers. comm.
<b>FmoA gene</b>		
F-Start-fmo	5'- ATG GCT CTT TTY GG -3'	Alexander et al. 2002
R-889-fmo	5'- CCG ACC ATN CCG TGR TG -3'	Alexander et al. 2002

<sup>a</sup> Numbering of the primers refers to the accordant position in the *E. coli* 16S rRNA gene sequence.

L1r binds to the end of the ITS region.

***PCR and cloning of the 16S-23S rRNA intergenic spacer (ITS) region***

The ITS region was amplified using standard PCR conditions and the universal ITS primers 1525f and L1r (Table 2, 3). For the BS1 culture suspected to contain different bacterial species, the primer GSB822f, which is specific for green sulfur bacteria (modified after (Overmann 1999)), was combined with L1r (Table 3). For sequencing, 1525f and L1r were combined (Table 3). PCR products were checked for double bands on 1.4 % agarose gels. Strains which yielded more than one operon were selected for cloning of the ITS. Those were *Chl. phaeobacteroides* UdG 6051, *Chl. limicola* DSM 246, *Chl. phaeobacteroides* DSM 266<sup>T</sup>, '*Cba. chlorovibrioides*' UdG 6043, *Cba. limnaeum* UdG 6045, *Ptc. sp.* 2K, *Pld. luteolum* E3P2, *Chl. phaeobacteroides* E3P3, *Chl. limicola* DSM 245<sup>T</sup>, *Chl. clathratiforme* PG, *Chl. limicola* UdG 6044, *Chl. phaeobacteroides* UdG 6046, *Chl. phaeobacteroides* Glu, *Cba. limnaeum* UdG 6040, *Cba. parvum* DSM 263<sup>T</sup>, *Chl. limicola* D1, *Chl. clathratiforme* DSM 5477<sup>T</sup> and the Black Sea bacterium BS-1 (Manske et al. 2005) (Suppl. Table 1). PCR products were ligated into the vector pCR2.1-TOPO (Invitrogen, Carlsbad, USA) and subsequently cloned into TOP10 chemically competent cells according to the instructions supplied by the manufacturer. Recombinant clones were obtained on LB agar plates supplemented with kanamycin (50 µg/ml), picked and grown in liquid LB medium with kanamycin (50 µg/ml). Plasmids were isolated and purified, randomly using the QIAprep Spin Miniprep Kit (Qiagen) or the Quantum Prep Plasmid Miniprep Kit (Bio-Rad, Hercules, USA). The ITS region was sequenced from the extracted plasmid DNA using the primers 1525f and L1r.

***Primer design and PCR for group 1  $\sigma^{70}$ -type sigma factor gene sequences (sigA)***

Sigma factor gene sequences for most of the GSB species were obtained with the primer pairs Sig208f/Sig827r and Sig312f/Sig827r, respectively (Table 3). The PCR reaction (Table 2) comprised 20 – 50 ng DNA, 1 µM of primer (0.5 µM for 312f), 0.2 mM of each dNTP, 5 µl of PCR buffer (10x, incl. 15 mM MgCl<sub>2</sub>) and 1 U of Taq polymerase (2 U for Sig312f) in a total volume of 50 µl (all chemicals by Perkin Elmer, Foster City, USA). Since the primer pairs Sig208f/Sig827r or Sig312f/Sig827r did not yield a *sigA* gene sequence in all investigated isolates, another primer pair was designed using an alignment of the available sequences and the *sigA* gene sequences of *Chl. tepidum* TLS and *Chlorobium chlorochromatii* CaD (Vogl et al. 2006). Conserved regions in the alignment were checked for their priming qualities, for self-complementaries, potential hairpin formation, dimer formation and palindromes, using the NetPrimer software (Premier Biosoft International, Palo

Alto, USA). Primers were then checked for specificity employing the option of *search for short, nearly exact matches* in the NCBI BLAST database ([www.ncbi.nlm.nih.gov/BLAST/](http://www.ncbi.nlm.nih.gov/BLAST/)). The production of a single amplification product in green sulfur bacteria was demonstrated by PCR for the two primers GSB-SigA-F4 and GSB-SigA-R1 (Table 3). For all subsequently analyzed green sulfur bacteria a pure amplification product was obtained with 50 ng of template DNA in standard PCR conditions (Table 2). The same primer pair was used for the sequencing reaction.

### ***BchG gene sequences***

*BchG* gene sequences were amplified using the published primer pair bchG-F/bchG-R (Table 3) with a modified cycling method (Table 2). For sequencing, a shortened reverse primer, BchG-R Seq, (Table 3) was used.

### ***FmoA gene sequences***

*FmoA* gene sequences were amplified using the published primer pair F-start-fmo/R-889-fmo (Table 3) with the conditions published by Alexander *et al.* (Alexander *et al.* 2002).

### ***Sequencing of PCR products***

Cycle sequencing was performed with the AmpliTaq FS BigDyeTerminator cycle sequencing kit (Applied Biosystems, Weiterstadt, Germany) following the protocol supplied by the manufacturer. Samples were analyzed on a capillary sequencer (ABI Prism377 DNA sequencer, Applied Biosystems).

### ***Phylogenetic trees***

Additionally to the sequences obtained during this study, a multitude of sequences of green sulfur bacterial strains and environmental sequences was withdrawn from the NCBI ([www.ncbi.nlm.nih.gov](http://www.ncbi.nlm.nih.gov)) and SILVA database ([www.arb-silva.com](http://www.arb-silva.com)) (Suppl. Table 1). For the 16S rRNA gene tree, the pre-aligned sequences of the SILVA database were checked for chimeras using Mallard (<http://www.ebi.ac.uk/Tools/msa/clustalw2>) (Ashelford *et al.* 2006). The sequences were phylogenetically analyzed using the ARB phylogeny software package (Ludwig *et al.* 1998). The sequences of the SILVA database were corrected manually based on 16S rRNA secondary structure information for *Escherichia coli*. From these sequences, phylogenetic trees were constructed employing the MAXIMUM LIKELIHOOD algorithm

(Fast DNA\_ML). The shorter sequences were inserted afterwards without changing the overall tree topology employing the PARSIMONY INTERACTIVE tool implemented in the ARB software package (Gich et al. 2001). No filter was used. Sequences obtained at the same sampling site with more than 99.7% sequence homology according to ARB neighbour joining were deleted from the tree since their difference is very likely due to mistakes of the taq polymerase used in the sequencing PCR.

For the ITS sequences, the *fmoA*, *bchG* and the *sigA* gene sequences the trees were calculated using the MAXIMUM LIKELIHOOD (Fast DNA\_ML) algorithm implemented in ARB (Ludwig et al. 1998).

### ***Concatenated phylogenetic tree***

For the concatenated analysis of three different genetic regions (*sigA*, *bchG* and *fmoA*) sequences of overall 32 strains were available, including type strains of all prominent clusters. The sequence lengths of all included organisms in the resulting alignments were customized and every third base was left out. The alignments were combined in ARB. Starting from this alignment, a concatenated tree was calculated using the MAXIMUM LIKELIHOOD (ML) algorithm implemented in ARB.

### ***Distance matrix comparisons (Mantel test)***

For the ITS region, the *bchG*, the *sigA* and the *fmoA* gene sequences of the GSB strains used for the concatenated tree, a matrix comparison analysis was conducted, comparing each of the distance matrices with the 16S rRNA gene distance matrix according to Mantel (Mantel 1967). Additionally, the first two bases as well as the third base of the protein coding genes were compared separately to the 16S rDNA and the ITS region to analyse the usually higher mutation rate of the non-coding third base.

For each marker, a distance matrix (symmetric dissimilarity matrix) was calculated using MEGA 5 (Tamura et al. 2011). Two matrices respectively were compared and plotted by using the ade4 library implemented in R (R Development Core Team, 2010). In this test, two matrices are multiplied, element-by-element, and the sum of these products is tested against the expected value of its quantity, the significance of which is evaluated by Monte Carlo techniques (Sokal 1979; Sokal *et al.* 1986). Since the number of element-by-element comparisons would be too high, a random sample of permutations (9999 permutations) is used. For more than 12 operational taxonomic units (OTU), a correlation greater than 0.5 will



be statistically significant. 100% correlation of permutational distribution would result in a correlation coefficient of 1.

### ***Coverage and diversity estimates***

The collectors curve was calculated using Analytic Rarefaction 1.3 by Steven M. Holland, available at <http://strata.uga.edu/software/index.html>. To estimate the diversity of green sulfur bacteria, clonal richness was calculated in EstimateS (Version 8.2.0) (Colwell 2005), using the Chao1 estimator (Chao 1984). Chao analysis was done without bias correction, using the classic formula. All calculations were performed on the species level with 97% sequence similarity and for identical sequences. In the latter case, sequences with a 99.7% similarity were considered to be identical to account for sequencing errors.

### ***Pair-wise sequence dissimilarity analysis and calculation of $k_a$ and $k_s$ values***

The  $k_a$  (the number of non-synonymous substitutions per non-synonymous site) and  $k_s$  (the number of synonymous substitutions per synonymous site) values were determined using DNaSP for all pair-wise comparisons of sequences (Librado and Rozas 2009).

The frequency distribution of (a) the pair-wise sequence dissimilarity values, as determined for the Mantel test, of the entire codon, only the first and second, or only third codon position, (b) the  $k_s$  values and (c) the  $k_a$  values were determined using the *density()* function with the default “gaussian” smoothing kernel as implemented in the free software environment R (Team 2010). The R package *multcomp* (Herberich et al. 2010) was utilized using tukey contrasts for multiple comparisons of means to determine significant differences between the value distributions among the three genes.

### ***Map of sampling sites***

A world map containing the sampling sites of the sequences used in the 16S rDNA tree was constructed. The locations of the sampling sites were gathered from the original literature, entries in the NCBI database (<http://www.ncbi.nlm.nih.gov/nucleotide>), or hand written documents of Norbert Pfennig in the archives of the DSMZ.

### ***Nucleotide accession numbers***

DNA sequences obtained in this study are available at the EMBL under accession numbers AM049271-AM049317, AM050117-AM050132, AM050061-AM050096, AM050351-AM050374 (Suppl. Table 1).

## **3.2 Experimental procedures to investigate the molecular basis of the interspecies interaction in “*C. aggregatum*”**

### ***Bacterial cultures and growth conditions***

Cultures of the phototrophic consortium "*C. aggregatum*" were grown in 10 l glass flasks containing anoxic K3 medium (pH 7.2) with 1 mM sulfide as electron donor and reductant (B. E. M. Kanzler et al. 2005) but lacking 2-oxoglutarate. The flasks were incubated at room temperature under continuous illumination of 25 mmol quanta m<sup>-2</sup> s<sup>-1</sup> incident light intensity. Under these conditions, the consortia form a dense biofilm on the inner surface of the culture vessel, which can be harvested separately from the accompanying bacteria (Pfannes *et al.* 2007). The pure culture of the epibiont *Chl. chlorochromatii* CaD was grown under the same conditions.

### ***Probes***

Biotin-labeled oligonucleotide probes and unlabeled helper probes were purchased from Eurofins MWG Operon (Ebersberg, Germany). GSB 822 was used as a specific probe for Green Sulfur Bacteria (Tuschak *et al.* 1999), whereas the ZS 207 probe (B. Kanzler 2004) was utilized to capture the central bacterial 16S rRNA. For magnetic capture, the probes were biotinylated at the 5'-end followed by a 30 thymine spacer separating the specific probes from the surface of the beads, thus increasing the RNA yield. To further increase the yield, unlabeled helper probes complementary to the consensus sequences upstream and downstream of GSB 822 and ZS 207 were used. GSB 822 helper probes were 15 bp long, whereas ZS 207 consisted of 18 bp upstream and 20 bp downstream. For fluorescence *in situ* hybridization cy3 labeled GSB 822 and ZS 442 (Kanzler 2004) probes were used.

***Nanoscale secondary ion mass spectrometry (NanoSIMS)***

To harvest “*Chlorochromatium aggregatum*” for NanoSIMS analysis, the consortia biofilm was carefully washed with K4 (K3-Medium buffered with 10 mM HEPES) (pH 7.2) medium lacking NaHCO<sub>3</sub>. The OD was measured at 650 nm and subsequent cultures were inoculated with an OD<sub>650</sub> of 1 and cultivated with an illumination of 25 mmol quanta m<sup>-2</sup> s<sup>-1</sup> incident light intensity. <sup>13</sup>C marked NaHCO<sub>3</sub> was added to a final concentration of 10 mM, at which point the 0 min control for this experiment was immediately taken and stopped with a final concentration of 2 % of paraformaldehyde. Further time points were taken at 15, 30, 45 and 60 minutes. The samples were left in paraformaldehyde overnight, washed twice with 1x PBS and either frozen as a 50 % glycerol stock or used immediately. The samples were filtered over a 0.2 µm polycarbonate filter (Millipore, Billerica, USA), washed twice with 1 x PBS and cut with a scalpel to fit onto a quarter of an ITO glass slide (Sigma-Aldrich, Germany, nr. 703184). The glass slide was cut to a square side length of 1.8 cm and marked with four different distinct marks in the glass surface for four different time points prior to the experiment. The filter was dried on the glass slide and then carefully removed leaving the consortia on the glass slide. Three to five consortia were selected for each time point paying attention to a clear visibility of the central bacteria within the epibionts. A series of pictures was taken for each of the consortia to later localize the consortia (63 to 10 magnification) corresponding to one of the marks in the glass slide in the NanoSIMS 50L.

Measurements were collected in both Image and Isotope mode on the CAMECA NanoSIMS 50L, with a resolving power of ~5,000. Four secondary ions were simultaneously collected: <sup>12</sup>C<sup>-</sup>, <sup>13</sup>C<sup>-</sup>, <sup>12</sup>C<sup>14</sup>N<sup>-</sup>, and <sup>12</sup>C<sup>15</sup>N<sup>-</sup>. In Image mode, a ~2.5 pA primary Cs<sup>+</sup>-beam with a nominal spot size of ~100-200 nm was used. The beam rastered over square regions of sides 5 to 20 µm depending on the size of the target, at 256 x 256 pixel resolution. A complete square raster or “frame” representing a layer of the target, was completed every 10 to 20 minutes, with several to 150 frames collected per target. Complete analysis of large targets (aggregates of diameter >15 µm) lasted up to 48 hours, but analysis of smaller targets (aggregates with diameter ~5 µm) were completed in approximately 12 hours. Measurements in Isotope mode were collected with an ~30-40 pA primary Cs<sup>+</sup>-beam over square regions of the same size but typically with a reduced resolution of 128 x 128 pixels, a per frame acquisition time of 30 seconds, and a total acquisition time of one to three hours. Although not yielding an image, Isotope mode measurements allowed more rapid screening of samples. Both Image and

Isotope modes were used to measure *E. coli* cell standards as described in Orphan Turk House 2009.

NanoSIMS images were processed with L'image (developed by L. Nittler, Carnegie Institution of Washington, Washington D.C.). Each series of frames was corrected for drift and detector dead time. When analyzing trends in the series of frames with depth, every 3-5 frames were binned into a single image. Discrete regions of Interest (ROI's), approximately 1  $\mu\text{m}$  in diameter, were hand-drawn on the ion images guided by the corresponding 3D FISH data and isotope ratios were subsequently calculated for the particular regions.

### ***Magnetic bead separation of $^{14}\text{C}$ - labeled RNA***

To specifically separate the 16S rRNA of the epibiont from the central bacterium, a consortia biofilm harvested in K3 Medium was mixed with 12.5 % ice-cold ethanol/phenol stop solution (5 % phenol pH 4.5–5.5 in 100 % ethanol) to avoid RNA degradation. Total RNA was isolated using phenol-chloroform (Chomczynski and Sacchi 1987), precipitated with 2.5 volumes 100 % ethanol and 1/10 Volume 3 M sodium acetate (pH 5.2) and diluted with dd  $\text{H}_2\text{O}$ . RNA concentrations were determined spectrophotometrically in a TECAN infinite M200 microplate reader (Tecan Austria GmbH, Grödig, Austria) at Ex/Em 490/520 using a 1:100,000 dilution of SYBR Green II RNA gel stain (Invitrogen Ltd, Paisley, UK) in 1x TBE buffer (pH 8.0). For magnetic capture, magnetic beads (Chemagen, M-PVA SAV1) were washed 3 times with twice the volume of 0.5x SSC (1x SSC is 0.15 M NaCl plus 0.015 M sodium citrate) using a magnetic-particle concentrator (Dynal MPC-S; Invitrogen) and afterward resuspended in the original volume of 0.5x SSC. For the epibiont RNA, 50  $\mu\text{l}$  and for the central bacterium 7  $\mu\text{l}$  beads were used, which were incubated with 700 pmol and 100 pmol 16S rRNA specific probes at room temperature in a thermomixer (thermomixer compact, Eppendorf, Hamburg) at maximum speed. After 30 min, the beads were washed 3 times with 0.5x SSC. For the hybridisation, beads labeled with GSB 822 probes were incubated with 5  $\mu\text{g}$  total RNA and 1.5 nmol of each helper probe, beads with ZS 207 probes with 10  $\mu\text{g}$  total RNA and 150 pmol of each helper probe in 100  $\mu\text{l}$  5x SSC. Prior to hybridisation, total RNA was incubated at 70°C for 5 min, placed immediately on ice and incubated for another 3 min. Hybridisation took place in a Mini-38 hybridisation oven (Bachofen Laboratoriumsgeräte, Reutlingen, Germany) under constant rotation at 68°C for the GSB 822 probe and 66°C for the ZS 207 probe for 30 min. Afterwards, the beads were washed 3 times with 0.5x SSC. During the second washing step, beads were incubated in 0.5x

SSC in a thermomixer for 3 min at 61°C for GSB 822 and 59°C for ZS 207 probes. The 16S rRNA was eluted by incubating the beads at 70°C for 3 min in 100 µl of DEPC treated ddH<sub>2</sub>O. The beads were collected by a magnetic-particle concentrator and the RNA was removed. The concentration was determined by a TECAN infinite M200 microplate reader using SYBR Green II. To test the specificity of the bead-capture, the eluted RNA was used for reverse transcription performed with Superscript III Reverse Transcriptase (Invitrogen GmbH, Darmstadt, Germany) using 50 ng of RNA and 200 ng of random hexamer primer (Eurofins MWG, Ebersberg, Germany). The cDNA obtained was used to amplify the 16S rRNA with the universal primers GC 8f and 1492r at an annealing temperature of 56°C. Amplification products were sequenced as described previously (Overmann *et al.* 1999) and separated on the basis of melting behavior by denaturing gradient gel electrophoresis (DGGE) using a gradient consisting of 30 % to 70 % denaturing agents (Muyzer *et al.* 1993).

The bead-capture technique was used to conduct a series of <sup>14</sup>C labeling experiments. Therefore, biofilms obtained from "*C. aggregatum*" cultures grown in 10 l glass flasks were harvested in K4 Medium (K3-Medium buffered with 10 mM HEPES) (pH 7,2) lacking NaHCO<sub>3</sub>. The OD<sub>650</sub> was adjusted to 1. [<sup>14</sup>C] sodium bicarbonate (50-60 mCi/mmol) (Hartmann Analytic GmbH, Braunschweig, Germany) was added to a concentration of 30 µM and the culture was distributed among gastight screw-cap test tubes. All test tubes were incubated on a rollerdrum (New Brunswick TC-7, New Brunswick Scientific, Edison, USA) under continuous illumination of 25 mmol quanta m<sup>-2</sup> s<sup>-1</sup> incident light intensity. Four different experiments were carried out: In the first experiment, samples were taken after 0 h, 2 h, 6 h, 12 h, 24 h, 48 h and 96 h. As a negative control, test tubes incubated in the dark were sampled after 0 h, 24 h and 96 h. From each test tube, 10 µl of the culture were filtered onto white polycarbonate filters of 0.2 µm pore size (Millipore, Billerica, USA) for total cell count and determination of consortia numbers via fluorescence *in situ* hybridization (see below). For subsequent pigment analysis, 2 ml of each tube were pelleted and frozen. After thawing however, the addition of 1 ml acetone led to cell lysis, so that the radioactivity of the whole cell extract was measured in a TriCarb 2100 TR liquid scintillation counter (Canberra-Packard GmbH, Schwadorf, Austria). In the second experiment, either 1 mM of 2-oxoglutarate or a mix of the three branched-chain amino acids (BCAA) valine, leucine and isoleucine each in a concentration of 1 mM were added to the medium in order to test whether externally added substrates have an influence on the incorporation of radiolabel into the two organisms. In the positive control, no supplements were used. The negative control was

incubated in the dark with 1 mM 2-oxoglutarate and 1 mM of each BCAA. Samples were taken after 0 h, 6 h, 24 h and 48 h and accordingly after 0 h, 24 h and 48 h for the negative control. For the third experiment, the 20 proteogenic amino acids were each added to a test tube in a concentration of 1 mM respectively. The positive control was treated as in experiment number 2. Samples were taken after 24 h. The experiment was repeated using 0.3 mM of each amino acid to avoid growth inhibition at higher concentrations and 1 mM molybdate to inhibit the activity of sulfate reducing contaminating bacteria in the enrichment culture (Yadav and Archer 1989). All experiments were carried out in three parallels. At each sampling, 5 ml of each test tube were mixed with 12.5 % ice-cold ethanol/phenol stop solution and the 16S rRNA of the epibiont and central bacterium were isolated as described above. The RNA concentration was determined by a TECAN infinite M200 microplate reader using SYBR Green II, the radioactivity was measured in a TriCarb 2100 TR liquid scintillation counter. The significance of the results was determined by a t-test analysis conducted in SigmaPlot (Systat Software Inc.).

### ***Cell counting***

The oligonucleotide probes GSB822 and ZS442 (B. Kanzler 2004) targeting the epibiont and the central bacterium of "*Chlorochromatium aggregatum*", were used for fluorescence *in situ* hybridisation (FISH). The probes were labeled with Cy3 (carboxymethylrhodamine) as the fluorescent dye. Cells were filtered onto white polycarbonate filters of 0.2 µm pore size (Millipore, Billerica, USA) and fixed with 4 % paraformaldehyde for 30 min at room temperature. Filters were cut into pieces and immersed in 300 µl hybridization buffer containing 900 mM NaCl, 20 mM Tris-HCl (pH 7.5), 20 % formamide, 1 % blocking solution (Roche Diagnostics GmbH, Mannheim, Germany), 0.01 % SDS and 167 nM of the Cy3-labelled probe. The samples were hybridized for 2 h at 46 °C followed by a washing step in hybridization buffer for 15 min at 48 °C. Filter sections were rinsed with sterile filtered dd H<sub>2</sub>O, dipped in 80 % ethanol for 1 min to reduce background staining and air dried on a microscope slide. After counterstaining with 4',6-diamidino-2-phenylindole (DAPI), cells were counted using epifluorescence microscopy with a Zeiss AxioStar Plus microscope equipped with a 100x oil immersion objective (Plan Apochromat 100x/1.4, Zeiss) and a 100-square counting grid (Zeiss).

***Fluorescence-Recovery-After-Photobleaching (FRAP)-analysis***

To test if free diffusion between the symbiotic partners of “*C. aggregatum*” is possible, a FRAP analysis with the soluble fluorescence dye calcein was performed. To perform the FRAP-analysis, 0.5 ml of a consortia biofilm resuspended in K3 medium to an OD<sub>650</sub> of 1 and mixed with 10 µl of calcein-AM (1 mg/ml in dimethylsulphoxide) (Invitrogen GmbH, Karlsruhe, Germany). The suspension was incubated at room temperature in the dark for 30 min. Cells were harvested by gentle centrifugation and resuspended three times in dye-free K3 medium. The suspension was then incubated in the dark for another 90 min prior to imaging. Cells were imaged with an upright Leica SP5 confocal laser scanning microscope using a 63.0x water-immersion objective and the 488 nm line of a 100 mW argon laser as the excitation source. The initial idea to bleach the central bacterium could not be carried out since the central bacterium was not stained by calcein-AM. Therefore, one of the epibionts was bleached for 0.5 seconds by switching the laser from 4% to 100% power using a 41.3 µm pinhole. Pictures series were recorded from one second prior to 30 seconds after bleaching.

***Sample preparation for isotopologics***

Biofilms obtained from “*C. aggregatum*” cultures were harvested in K4 Medium (pH 7.2) lacking NaHCO<sub>3</sub> and the OD<sub>650</sub> was adjusted to 1. Overall, four experiments were conducted. Firstly, consortia were incubated in K3 medium with 10 mM NaH<sup>13</sup>CO<sub>3</sub> (Sigma-Aldrich, Germany) in the light to investigate the incorporation of <sup>13</sup>C-label derived from H<sup>13</sup>CO<sub>3</sub><sup>-</sup> into the amino acids of the epibiont and the central bacterium and as a positive control conducted under standard cultivation conditions. Samples were taken after 1 h and 3 h. Secondly, to investigate if the epibiont in the free-living state shows different incorporation characteristics than in the associated state, an epibiont pure culture was incubated in a second essay under the conditions of the first one. In the third essay, consortia were cultivated in the light in K3 medium with either 1 mM [U-<sup>13</sup>C] 2-oxoglutarate or 1 mM [U-<sup>13</sup>C] leucine (Cambridge Isotope Laboratories Inc., Andover, USA) to investigate the effect of external substrates under standard cultivation conditions as opposed to the incubation of consortia with 30 µM NaHCO<sub>3</sub> used in the <sup>14</sup>C-labeling experiments. Samples were taken after 3 h. In the fourth essay, consortia were cultivated in the dark in K3 medium lacking NaHCO<sub>3</sub> with the addition of either 1 mM [U-<sup>13</sup>C] 2-oxoglutarate or 1 mM [U-<sup>13</sup>C] leucine to investigate if either one of these substrates can be heterotrophically incorporated with an inactive epibiont. Samples were taken after 3 h and all experiments were carried out in two parallels.

After incubation, the cultures were pelleted and resuspended in K4 medium lacking H<sub>2</sub>S and NaHCO<sub>3</sub> to an OD<sub>650</sub> of 4. To disaggregate the consortia into single cells, they were incubated in a water bath at 68°C for 10 min. During the incubation, aggregates were disrupted by passing the culture through a syringe (0.80 mm x 120 mm) several times. Subsequently, the central bacterial cells were separated from the epibiont by CsCl equilibrium density centrifugation. The two distinct bands of the epibiont and the central bacterium were removed with a syringe and needle (0.80 mm x 120 mm) and 20 µl were removed for FISH analysis, respectively, to determine the contamination by other bacteria in each band. The remaining volume was mixed with two times sterile ddH<sub>2</sub>O, pelleted and shock frozen in liquid nitrogen.

#### ***Hydrolysis of cellular protein and silylation of the liberated amino acids***

Approximately 10 mg of lyophilized cells were suspended in 1.5 ml of 6 M hydrochloric acid. The mixture was heated at 105°C for 24 hours under an inert atmosphere. The hydrolyzate was dried under a stream of nitrogen at 70°C. The residue was suspended in 200 µl of concentrated acetic acid and placed on a column of Dowex 50W×8 (H<sup>+</sup> form, 200-400 mesh, 5 × 10 mm). The column was washed twice with 1 ml of water and is then developed with 1 ml of 4 M ammonium hydroxide. An aliquot of the eluate (200 µl) was dried under a stream of nitrogen and the residue is dissolved in 50 µl of dry acetonitrile. 50 µl of N-(tert-butyl-dimethylsilyl)-N-methyl-trifluoroacetamide containing 1% tert-butyl-dimethylsilylchlorid (Sigma-Aldrich) was added. The mixture was kept at 70°C for 30 min. The resulting mixture of tert-butyl-dimethyl-silyl derivatives (TBDMS) of amino acids was used for GC/MS analysis without further work-up.

#### ***Gas chromatography - mass spectrometry (GC/MS)***

GC/MS analysis was performed with a GC-QP2010 Plus instrument (Shimadzu, Duisburg, Germany) equipped with a fused silica capillary column (Equity TM-5; 30 m × 0.25 mm, 0.25 µm film thickness; SUPELCO, Bellafonte, PA) and the mass detector working with electron impact ionization at 70 eV. An aliquot (3 µl) of a solution containing TBDMS amino acids was injected in 1:10 split mode at an interface temperature of 260°C and a helium inlet pressure of 75 kPa. The column was developed at 150°C for 3 min and then with a temperature gradient of 7°C/min to a final temperature of 280°C that was held for 3 min. Data were collected using the LabSolutions software (Shimadzu, Duisburg, Germany). Selected ion



monitoring experiments were carried out on the fragments summarized in Table 4 using a 0.3 s sampling rate. Samples were analysed at least three times. Data were processed as described before (Lee *et al.* 1991) affording the molar excess of carbon isotopomer groups (M+1, M+2, ..., M+n, with M being the mass ion of the fragment under study and n being the number of C-atoms of the respective fragment). In the case of leucine and isoleucine, a five carbon atom fragment is analysed. Therefore, the two amino acids are indicated with M+5. Excess of multiple labelled isotopologues was calculated according to the equation,  $[2 * (M + 2) + 3 * (M + 3) + \dots + n * (M + n)] / n$ .

This procedure was required since natural carbon already carries  $^{13}\text{C}$  at an abundance of 1.10 %. Isotopes of silicon also contribute to the detected mass fractions of the TBDMS-derivatives and their fragments with  $^{29}\text{Si}$  at 4.67 % and  $^{30}\text{Si}$  at 3.10 % abundance, respectively. These natural contributions of stable isotopes had to be subtracted from the overall mass fractions that were detected. Since multiple labeling can occur for statistical reasons, the natural contribution of multiple labeled isotopologues was calculated and subtracted from the original mass data (for more details, see also Pickup and McPherson, 1976).

**Table 4.** Analyzed fragments of TBDMS amino acids in the order of their retention time (RT). ‘\*’ indicate fragments with higher error bars according to Antoniewicz *et al.*, (2007).

Amino acid	RT [min]	(M - 57) <sup>+</sup>	(M - 85) <sup>+</sup>
Ala	6.7	260	232
Gly	7.0	246	218
Val	8.4	288	260
Leu	9.0		274
Ile	9.5		274
Pro	10.0	286*	258*
Ser	13.2	390	362
Thr	13.6	404	376
Phe	14.6	336	308
Asp	15.4	418	390
Glu	16.9	432	404

### **Genome analysis**

To enrich the central bacteria, several rounds of CsCl equilibrium density gradient centrifugation were performed as described above. The DNA from these central bacterium enrichments was prepared as previously described (Jogler *et al.* 2009; Zhou *et al.* 1996). Briefly, 675 µl of DNA extraction buffer was used to resuspend the individual cell pellets from the CsCl density gradient centrifugation. A 2.5-µl volume of proteinase K (20 mg/ml) was added, and samples were incubated at 37°C with agitation (220 cycles/min). Afterward, 75 µl of 20 % sodium dodecyl sulfate was added, and samples were heated to 65°C for 2 h. Chloroform and isoamyl alcohol (1:24; 750 µl) were added, and the suspension was incubated at RT for 5 min, with the tube was inverted every minute. After centrifugation at  $6,000 \times g$  for 20 min at RT, the aqueous supernatant was transferred into a fresh 1.5-ml reaction tube and the DNA was precipitated with 350 µl of isopropanol at RT for 1 h. The DNA pellet was obtained by centrifugation for 20 min at RT and was washed twice with 70 % ethanol at 4°C. Finally, the DNA was resuspended in water by incubation at 55°C for 2 h. The DNA was frozen at -20°C and sent to the laboratory of Prof. Dr. Donald A. Bryant at the Department of Biochemistry and Molecular Biology of the Pennsylvania State University, USA for sequencing and processing of the data.

### **In silico subtractive hybridization**

In order to investigate if the epibiont genome shows signs symbiosis specific alteration like the reduction of genome size or a change in the G/C-content, it was compared to the other eleven sequenced green sulfur bacterial genomes. The comparison was conducted with the Phylogenetic Profiler available at the DOE Joint genome Institute website (<http://img.jgi.doe.gov>). The genome of *Chl. chlorochromatii* CaD ([http://genome.jgi-psf.org/finished\\_microbes/chlag/chlag.home.html](http://genome.jgi-psf.org/finished_microbes/chlag/chlag.home.html)) was screened for single genes which had no homologs based on BLASTP alignments against the other 11 publicly available genome sequences of the green sulfur bacteria *Chlorobaculum parvum* NCIB 8327, *Cba. tepidum* ATCC 49652<sup>T</sup>, *Chl. ferrooxidans* DSM 13031<sup>T</sup>, *Chl. limicola* DSM 245<sup>T</sup>, *Chl. luteolum* DSM 273<sup>T</sup>, *Chl. phaeobacteroides* BS1, *Chl. phaeobacteroides* DSM 266<sup>T</sup>, *Chl. phaeovibrioides* DSM 265, *Chl. clathratiforme* DSM 5477<sup>T</sup>, *Chloroherpeton thalassium* ATCC 35110<sup>T</sup> and *Prosthecochloris aestuarii* DSM 271<sup>T</sup> (<http://img.jgi.doe.gov/cgi-bin/geba/main.cgi?section=TaxonList&page=taxonListPhylo&pidt=14955.1250667420>). A

maximum e-value of  $10^{-5}$  and a minimum identity of 30% were applied for identification of homologs.

### ***Transcriptome analysis***

The consortia biofilm of a 10 l glass flask was harvested in K3 medium and the OD<sub>650</sub> was adjusted to 1. One half of the culture was incubated in the light, the other half in the dark with the addition of 1 mM valine, 1 mM leucine, 1 mM isoleucine and 1 mM 2-oxoglutarate to investigate if the central bacterium is metabolically active with external substrates but an inactive epibiont. In addition, the transcriptome of *Chl. chlorochromatii* CaD3<sup>T</sup> was analyzed to improve the coverage of the free-living epibiont transcriptome recently acquired (Wenter *et al.* 2010). Therefore, an epibiont pure culture was incubated in 500 ml K3 Medium in the light. After 3 days, total RNA of the three cultures was isolated as described above and purified with the RNeasy MinElute Cleanup Kit (Qiagen, Hilden, Germany) according to the protocol of the manufacturer. The RNA was treated with the *TURBO DNA-free*<sup>TM</sup> Kit (Ambion, Austin, TX, USA) to remove DNA contamination. RNA quality was analyzed using an Agilent 2100 bioanalyser with a RNA LabChip (Agilent Technologies, Santa Clara, USA). The separation of mRNA from 10 µg of total RNA was performed with the *MICROBExpress*<sup>TM</sup> Kit (Ambion) for bacterial mRNA enrichment according to the instructions of the manufacturer. The efficiency of rRNA depletion in the mRNA extract was analyzed using an Agilent 2100 bioanalyser. The mRNA was transcribed into double-strand cDNA and amplified using the Ovation RNA-Seq System (NuGEN Technologies, Bembel, Netherlands) according to the protocol of the manufacturer starting with 30 ng of enriched mRNA. Amplified cDNA was purified using the Zymo Clean and Concentrator Kit (Zymo Research Corporation, Irvine, USA). The cDNA samples were further prepared and sequenced using 36 bp single end sequencing on a Genome Analyzer II x (Illumina). Libraries of 300 bp were prepared according the manufacturer's instructions "Preparing Samples for Sequencing of mRNA". Cluster generation was performed using the Illumina cluster station, sequencing on the Genome Analyzer IIx followed a standard protocol. The fluorescent images were processed to sequences using the Genome Analyzer Pipeline Analysis software 1.7 (Illumina). The sequence output (36 single end short reads) of the Genome Analyzer IIx was transformed to FastQ format.

Comparison of transcript levels both within and between samples was facilitated by normalizing for RNA length and total read number by applying the reads per kilobase per

million mapped reads (RPKM) model (Mortazavi *et al.* 2008) applying the software CLC Genomics Workbench (<http://www.clcbio.com>). Analogous to analyzing differential expression in microarray essays, genes were considered non-differentially expressed when the relation of their RPKM-values was smaller than 0.5 or bigger than 2 (Bullard *et al.* 2010). Since we estimate that for each central bacterium, twenty epibionts exist in the consortia enrichment culture, the RPKM values for the free-living and symbiotic *Chl. chlorochromatii* were divided by 20 in order to increase the comparability of the transcriptome analysis of the two organisms.

### ***HPLC analysis of amino acid excretion***

To analyze the excretion of amino acids, filtrates (0.2  $\mu\text{m}$ ) of enrichment and epibiont pure culture supernatants were frozen in liquid nitrogen, concentrated by lyophilization and stored at  $-80^{\circ}\text{C}$ . Prior to HPLC, the lyophilisate was dissolved in distilled water ( $0.21 \text{ mg} \cdot \text{ml}^{-1}$ ) and filtered through 0.22  $\mu\text{m}$  pore size acrodisc syringe filters with low protein binding affinity (Pall Corporation, East Hills, USA). Concentrations of dissolved free amino acids (DFAA) were determined directly after ortho-phthaldialdehyde (OPA) derivatization (Lindroth and Mopper 1979) using the protocol of Grossart *et al.* (Grossart *et al.* 2007), and a HPLC system (Agilent 2000, Agilent, Santa Clara, USA) equipped with a C18 column (Waters). Dissolved combined amino acids (DCAA) were hydrolyzed with 6 mol HCl at  $155^{\circ}\text{C}$  for 1 h and then analyzed as DFAA.

## References

- Alexander, B., Andersen, J. H., Cox, R. P., and Imhoff, J. F. (2002). Phylogeny of green sulfur bacteria on the basis of gene sequences of 16S rRNA and of the Fenna-Matthews-Olson protein. *Archives of microbiology* 178(2), 131-140.
- Antoniewicz, M. R., Kelleher, J. K., and Stephanopoulos, G. (2007). Elementary metabolite units (EMU): a novel framework for modeling isotopic distributions. *Metab Eng* 9(1), 68-86.
- Ashelford, K. E., Chuzhanova, N. A., Fry, J. C., Jones, A. J., and Weightman, A. J. (2006). New screening software shows that most recent large 16S rRNA gene clone libraries contain chimeras. *Appl Environ Microbiol* 72(9), 5734-41.
- Bullard, J. H., Purdom, E., Hansen, K. D., and Dudoit, S. (2010). Evaluation of statistical methods for normalization and differential expression in mRNA-Seq experiments. *BMC Bioinformatics* 11, 94.
- Chao, A. (1984). Non-parametric estimation of the number of classes in a population. *Scandinavian Journal of Statistics* 11, 265-270.
- Chromczynski, P., and Sacchi, N. (1987). Single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction. *Anal. Biochem.* 162, 156-159.
- Coolen, M. J. L., and Overmann, J. (2000). Functional exoenzymes as indicators of metabolically active bacteria in 124,000-year-old sapropel layers of the eastern Mediterranean Sea. *Applied and environmental microbiology* 66(6), 2589-2598.
- Gich, F., Garcia-Gil, J., and Overmann, J. (2001). Previously unknown and phylogenetically diverse members of the green nonsulfur bacteria are indigenous to freshwater lakes. *Arch Microbiol* 177(1), 1-10.
- Grossart, H. P., Tang, K. W., Kiorboe, T., and Ploug, H. (2007). Comparison of cell-specific activity between free-living and attached bacteria using isolates and natural assemblages. *FEMS Microbiol Lett* 266(2), 194-200.
- Herberich, E., Sikorski, J., and Hothorn, T. (2010). A robust procedure for comparing multiple means under heteroscedasticity in unbalanced designs. *PLoS One* 5(3), e9788.
- Jogler, C., Lin, W., Meyerdierks, A., Kube, M., Katzmann, E., Flies, C., Pan, Y., Amann, R., Reinhardt, R., and Schuler, D. (2009). Toward cloning of the magnetotactic metagenome: identification of magnetosome island gene clusters in uncultivated

- magnetotactic bacteria from different aquatic sediments. *Appl Environ Microbiol* 75(12), 3972-9.
- Kanzler, B. (2004). Diploma thesis, Ludwig Maximilians University, Munich. p. 87
- Kanzler, B. E. M., Pfannes, K. R., Vogl, K., and Overmann, J. (2005). Molecular characterization of the nonphotosynthetic partner bacterium in the consortium "Chlorochromatium aggregatum". *Applied and environmental microbiology* 71(11), 7434-7441.
- Lee, W. N., Byerley, L. O., Bergner, E. A., and Edmond, J. (1991). Mass isotopomer analysis: theoretical and practical considerations. *Biol Mass Spectrom* 20(8), 451-8.
- Librado, P., and Rozas, J. (2009). DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. *Bioinformatics* 25(11), 1451-2.
- Lindroth, P., and Mopper, K. (1979). High performance liquid chromatographic determination of subpicomole amounts of amino acids by precolumn fluorescence derivatization with o-phthalaldehyde. *Analytical Chemistry* 51, 1667-1674.
- Ludwig, W., Strunk, O., Klugbauer, S., Klugbauer, N., Weizenegger, M., Neumaier, J., Bachleitner, M., and Schleifer, K. H. (1998). Bacterial phylogeny based on comparative sequence analysis. *Electrophoresis* 19(4), 554-568.
- Manske, A. K., Glaeser, J., Kuypers, M. M. M., and Overmann, J. (2005). Physiology and phylogeny of green sulfur bacteria forming a monospecific phototrophic assemblage at a depth of 100 meters in the Black Sea. *Applied and environmental microbiology* 71(12), 8049-8060.
- Mantel, N. (1967). The detection of disease clustering and a generalized regression approach. *Cancer Res* 27(2), 209-20.
- Mortazavi, A., Williams, B. A., McCue, K., Schaeffer, L., and Wold, B. (2008). Mapping and quantifying mammalian transcriptomes by RNA-Seq. *Nat Methods* 5(7), 621-8.
- Muyzer, G., De Waal, E. C., and Uitterlinden, A. G. (1993). Profiling of complex microbial populations by denaturing gradient gel electrophoresis analysis of polymerase chain reaction-amplified genes coding for 16S rRNA. *Applied and environmental microbiology* 59(3), 695-700.
- Overmann, J. (1999). Mahoney lake: A case study of the ecological significance of phototrophic sulfur bacteria, Vol. 15, pp. 251-288.

- Overmann, J., Coolen, M. J. L., and Tuschak, C. (1999). Specific detection of different phylogenetic groups of chemocline bacteria based on PCR and denaturing gradient gel electrophoresis of 16S rRNA gene fragments. *Archives of microbiology* 172(2), 83-94.
- Pfannes, K. R., Vogl, K., and Overmann, J. (2007). Heterotrophic symbionts of phototrophic consortia: members of a novel diverse cluster of Betaproteobacteria characterized by a tandem *rrn* operon structure. *Environmental microbiology* 9(11), 2782-2794.
- Sokal, R. R. (1979). Testing statistical significance of geographic variation patterns. *Systematic Zoology* 28, 227-231.
- Sokal, R. R., Smouse, P.E., Neel J.V. (1986). The genetic structure of a tribal population, the Yanomama Indians. XV. Patterns inferred by autocorrelation analysis. *Genetics* 114, 259-287.
- Tamura, K., Peterson, D., Peterson, N., Stecher, G., Nei, M., and Kumar, S. (2011). MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Mol Biol Evol* 28(10), 2731-9.
- R Core Development Team (2010) R: A language and environment for statistical computing. Vienna: R Foundation for Statistical Computing. Available: <http://www.R-project.org/>
- Tuschak, C., Glaeser, J., and Overmann, J. (1999). Specific detection of green sulfur bacteria by in situ hybridization with a fluorescently labeled oligonucleotide probe. *Archives of microbiology* 171(4), 265-272.
- Vogl, K., Glaeser, J., Pfannes, K. R., Wanner, G., and Overmann, J. (2006). *Chlorobium chlorochromatii* sp. nov., a symbiotic green sulfur bacterium isolated from the phototrophic consortium "Chlorochromatium aggregatum". *Archives of microbiology* 185(5), 363-372.
- Wenter, R., H • tz, K., Dibbern, D., Li, T., Reisinger, V., Ploscher, M., Eichacker, L., Eddie, B., Hanson, T., Bryant, D. A., and Overmann, J. (2010). Expression-based identification of genetic determinants of the bacterial symbiosis 'Chlorochromatium aggregatum'. *Environmental microbiology*.
- Widdel, F., Kohring, G. W., and Mayer, F. (1983). Studies on dissimilatory sulfate-reducing bacteria that decompose fatty acids - III. Characterization of the filamentous gliding *Desulfonema limicola* gen. nov. sp. nov., and *Desulfonema magnum* sp. nov. *Archives of microbiology* 134(4), 286-294.
- Yadav, V. K., and Archer, D. B. (1989). Sodium molybdate inhibits sulphate reduction in the anaerobic treatment of high-sulphate molasses wastewater. *Applied Microbiology and Biotechnology* 31(1), 103-106.

Zhou, J., Bruns, M. A., and Tiedje, J. M. (1996). DNA recovery from soils of diverse composition. *Appl Environ Microbiol* 62(2), 316-22.





## Chapter 4

### Results

#### 4.1 Results of the phylogenetic analysis

##### *16S rRNA gene sequence tree of GSB*

A phylogenetic tree comprising all sequences of the family *Chlorobiaceae* was constructed using the ARB software package. Of the 905 sequences retrieved from the SILVA database, 78 were removed after checking the sequences for chimeras using Mallard, 504 environmental sequences were removed because of their homology to already existing sequences from the same habitat. Of the remaining 323 green sulfur bacterial 16S rRNA gene sequences, a phylogenetic tree was calculated (Suppl. Fig. 1). The tree comprises the four basic clusters proposed by Imhoff (2003) and their subclusters 2a/b, 3a/b and 4a/b. With 96 sequences, group 3 represents the biggest cluster (3b: 53; 3a: 43), followed by 91 sequences in group 4 (4a: 46; 4b: 45). Group 2 in comparison consists of only 45 sequences (2a: 11; 2b: 34), group 1, the marine group, of 80. 11 sequences, including *Chloroherpeton thalassium* could not be assigned to any of the four clusters.

The most cultivated GSB are found in the marine group, with 41 isolated or enriched cultures, followed by 34 in group 4 (4a: 14; 4b: 20). Only nine GSB were cultivated from group 2 (2a: 3; 2b: 6) and 20 from group 3 (3a: 7; 3b: 13). Thus, a total of 106 GSB were isolated or enriched so far.

Two of the 16S rDNA sequences were ubiquitously distributed among the globe. The first sequence of group 4 comprises the cultivated strains *Chlorobaculum limnaeum* DSM 1677<sup>T</sup> (Lake Kinneret, Israel), 1549 (Solling, Germany), UdG 6040 (La Puda, Spain), UdG 6042 (Lake Vilar, Spain), C (unknown) and the environmental sequences DQ401523 (National University of Ireland), DQ383307 (photobioreactor), FJ484103 (El Zacaton, Mexico), DQ191736 (anaerobic wastewater bioreactor, Ireland), AY394842 (Yellowstone National Park, USA) and EU478457 (contaminated sediment). The second ubiquitous sequence is found in group 1 consisting of *Prosthecochloris vibrioformis* UdG7006Lms (Massona Lagoon, (Emporde Marshes), Spain), EP2403 (Prevost Lagoon, France), CHP 3402

(Chiprana Lake, Spain), DSM 260<sup>T</sup> (Moss Landing, California, USA), *Prosthecochloris phaeum* CIB 2401 (Mallorca, Balears Island, Spain) and *Prosthecochloris* sp. GSB-2 (Khakassia, salt Lake Shunet, Russia).

Interestingly, with the exception of one environmental sequence, all GSB derived 16S rDNA sequences collected from corals are clustering together. In this cluster of 8 sequences within the marine group, *Prosthecochloris* 3M is the only cultivated strain. All of the sequences have 97% sequence similarity to each other, with the exception of environmental sequence AY038478 which has only 97% sequence similarity to four of the sequences. Thus, it seems as if one species of GSB has specialized on living in an association with corals.

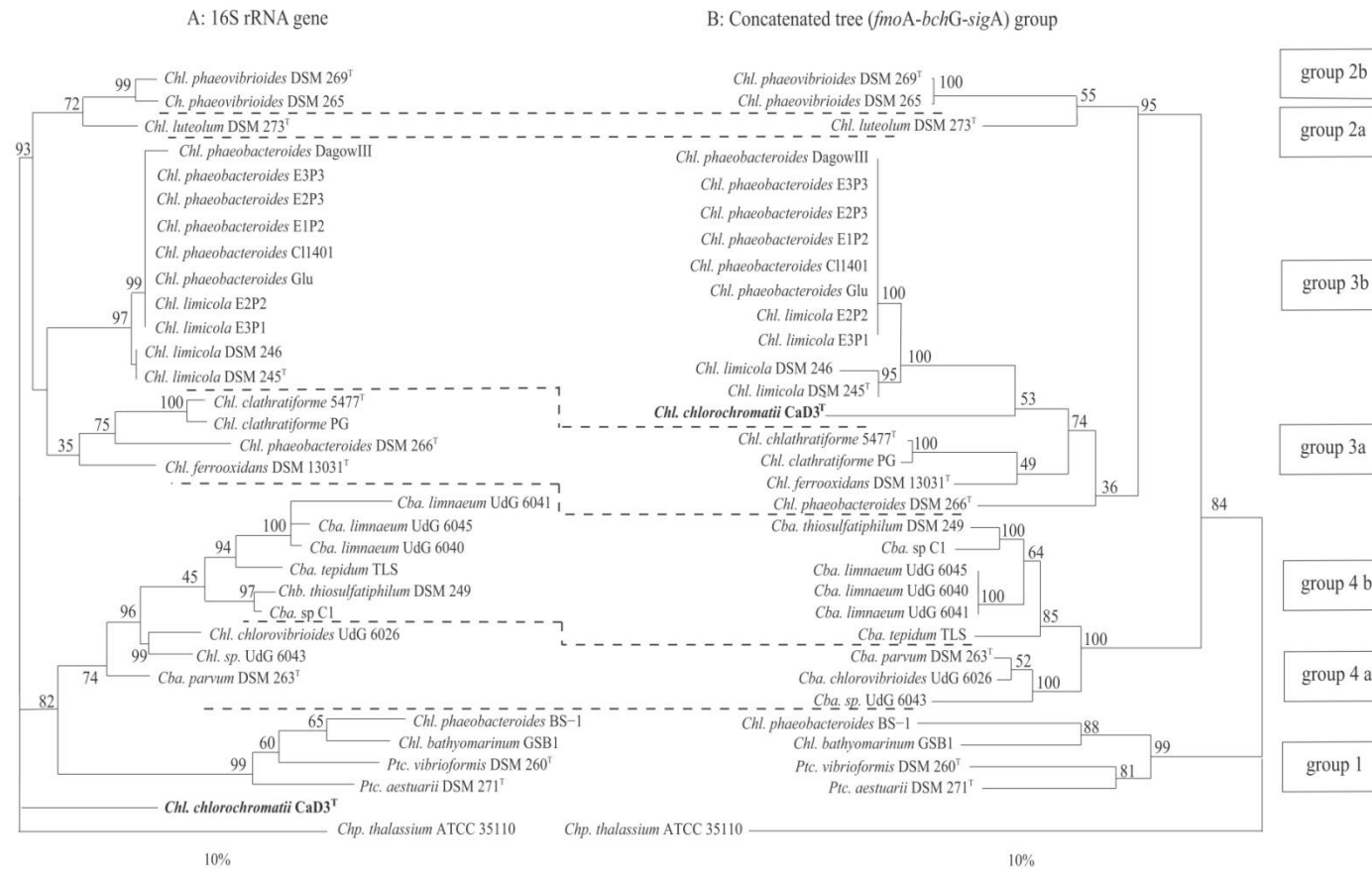
The majority of the sequences derived from warm or hot habitats are found in group 4. This includes all sequences of GSB found in thermal springs of Yellowstone National Park, USA, the mesophilic strain *Chlorobaculum thiosulfatophilum* DSM 249<sup>T</sup> from Tassajara Hot Spring, USA, an environmental sequence from a submarine hot spring in Japan, one from Thermopiles hot springs in Greece and the thermophilic species *Chlorobaculum tepidum* ATCC 49652<sup>T</sup>. But strains isolated from hot environments have also been found in group 1, *Chlorobium bathyomarinum* GSB1 from a thermal vent at the Juan de Fuca Ridge in the Pacific Ocean as well as *Prosthecochloris* sp. 5H2 also from the submarine hot spring in Japan and in group 3b, *Chlorobium limicola* DSM 245<sup>T</sup> from Gilroy Hot Spring, USA.

### ***Comparison of the concatenated tree with the 16S rDNA gene tree***

Of the three marker genes *fmoA*, *bchG* and *sigA*, a concatenated tree leaving out the third, non-coding base, was calculated (Fig. 3). Phylogenetic distances were longer in the concatenated tree than in the 16S rDNA tree, but the tree topology was nearly identical to that of the 16S rDNA tree. The groups and subgroups proposed by Imhoff were equally reflected in the concatenated and the 16S tree. Only the concatenated sequence of *Chl. chlorochromatii* CaD, the epibiont of the phototrophic consortium “*Chlorochromatium aggregatum*” clustered differently in the concatenated than in the 16S rDNA tree. While the 16S rDNA sequence cannot be assigned to any of the four groups, the concatenated sequence falls within group 3a of the concatenated tree.

Strikingly, the strains *Chl. phaeobacteroides* E1P2, E2P3, E3P3, C11401, DagowIII, Glu and *Chl. limicola* E2P2, E3P1 have identical concatenated sequences. With the exception of *Chl. phaeobacteroides* DagowIII, the 16S rDNA sequences of these strains are identical as well. Another group of identical concatenated sequences comprises the strains *Cba. limnaeum*

UdG 6045, 6040 and 6041. However, the corresponding 16S rRNA gene sequences differ considerably. This however is easily explained by looking at the 16S rRNA gene sequences. These contain several 'N's where the correct nucleotides could not be assigned, which speaks for an overall flawed sequencing of these strains.



**Figure 3.** Phylogenetic tree calculated from concatenated sequences of the coding bases of *fmoA*, *bchG* and *sigA* gene sequences in comparison with 16S rRNA tree. Trees were calculated with ML (FAST\_DNAML) using the ARB software package. Naming of the groups refers to Imhoff (2003). Bars denote fixed substitutions per nucleotide.

### ***Distance matrix comparisons (Mantel test)***

To investigate how the three marker genes *bchG*, *fmoA*, *sigA* and the ITS correlate with the 16S rDNA, a distance matrix comparison according to Mantel was conducted. Additionally to the complete sequences, a Mantel test for the third base of each codon and the first two bases of each codon were separately calculated. The *fmoA* gene sequence showed the highest correlation to the 16S rDNA (Table 5), followed by *bchG*. Only a slight correlation was found between the *sigA* gene sequence and the 16S rDNA. For the ITS, no correlation could be assessed.

**Table 5.** Distance matrix comparison of *sigA*, *bchG*, *fmoA* and the ITS region against the 16S rDNA

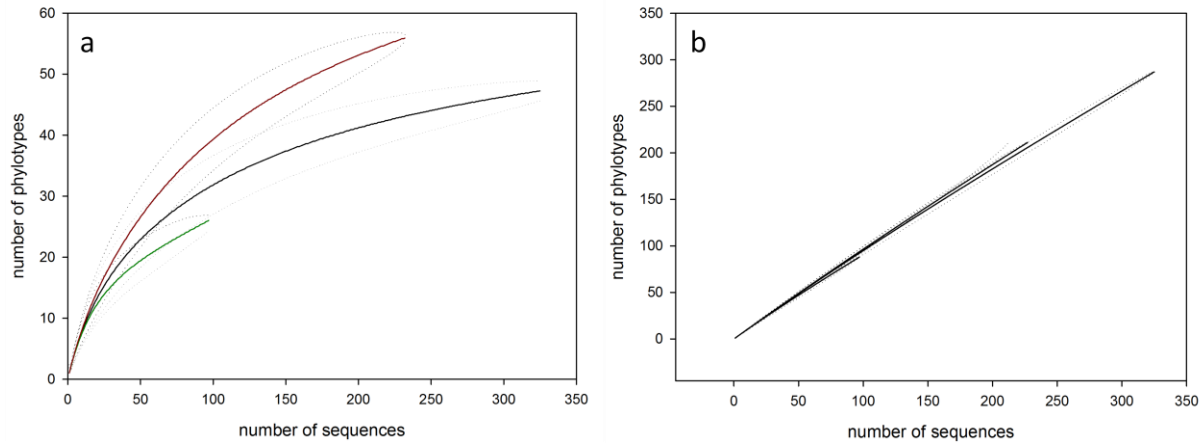
	3. base	1. and 2. base	all 3 bases
<i>sigA</i>	0.50	0.69	0.59
<i>bchG</i>	0.55	0.85	0.66
<i>fmoA</i>	0.64	0.86	0.77
ITS	-	-	0.49

As expected, the first two and thus coding bases of the protein coding genes are stronger correlated to the 16S rDNA than the complete sequence. In addition, the ratio of correlation between the three genes is shifted. With a coefficient of 0.85, *bchG* correlates nearly as much as *fmoA* (0.86) to the 16S rDNA. The reason for this shift lies within the third, non-coding base, which correlates stronger in *fmoA* than in *bchG*, which causes an overall higher correlation of the complete *fmoA* sequence to the 16S rDNA. The third base of the *bchG* gene sequence is hardly correlated and for *sigA*, no correlation is found.

### ***Coverage and diversity estimates***

To calculate how much of the *Chlorobiaceae* biodiversity is covered by the presented sequence analysis, a rarefaction (Fig. 4. a, b) and species richness analysis was conducted. The rarefaction analyses conducted with 97% sequence similarity using cultivated and uncultivated sequences already reached the saturation stage with a value of 47.0 phylotypes (Fig. 4. a). This is close to the Chao1 estimation of the maximum number of phylotypes of 56.6 (95% confidence interval 51.7 and 75.4). Interestingly, the slope of the rarefaction analysis for the cultivated strains ascends less steep than the slope of all sequences combined. For the Chao1 analysis, a maximum number of phylotypes of only 39.8 (95% confidence

interval 29.4 and 81.6) is estimated. Since the slope of the cultivated strains is less steep than the overall slope, the curve of the uncultivated sequences shows a steeper ascend with a Chao1 estimation of 70.3 (95% confidence interval: 60.8 and 98.7).



**Figure 4.** Rarefaction analysis of all available green sulfur bacterial sequences (—), sequences from cultivated strains (—) and environmental sequences (—) using a 97.0% (a) and a 99.7% (b) phylotype similarity. A confidence interval of 95% is given (···).

In the case of sequence similarity, for which of 99.7% similarity were considered to be identical to account for sequencing errors, uncultivated, cultivated and both sequences together do not reach saturation. The Chao1 estimation for sequences with less than 99.7% sequence similarity is 1222 (95% confidence interval 521 and 3054) for the cultivated strains, 2423 (95% confidence interval 1357 and 4481) for the uncultivated sequences and 3491 (95% confidence interval 2046.4 and 6122) for all sequences combined.

Thus, on the species level of 97% sequence similarity, the presented study comprises nearly the whole diversity of green sulfur bacteria. However, on the 99.7% sequence similarity level, the majority of the sequences has yet to be discovered.

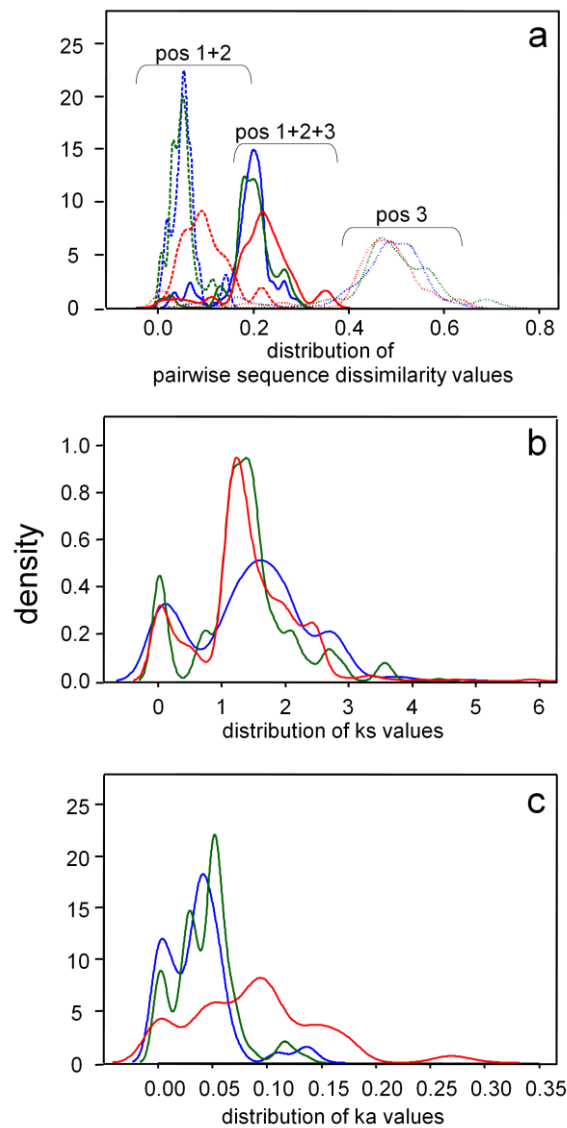
#### ***Pair-wise sequence dissimilarity analysis and calculation of $k_a$ and $k_s$ values***

To investigate which of the three marker genes provides the highest resolution when resolving the phylogeny of the GSB, the pair-wise sequence dissimilarity of each marker gene was calculated. Again, the complete sequence, the first two bases of each codon and the third base were analysed separately. For the third base, the pair-wise sequence dissimilarity between the three marker genes is nearly identical (Fig. 5. a). The differences are not significant. Thus, the

third base evolved similarly in all three marker genes leading to a comparable amount of nucleotide differences in the three marker genes. For the coding bases, the pair-wise sequence dissimilarity for *fmoA* is significantly higher than that of the two other marker genes, which behave identical in this analysis. Although this difference diminishes when the third base is taken into account, it is still significant when analysing the complete sequence of the marker genes. Therefore, the *fmoA* sequence not only shows the highest correlation to the 16S rDNA according to the Mantel test, but also provides the highest phylogenetic resolution of the three marker genes investigated since the most differences in the 32 sequences occur in this gene.

To find out at which positions the differences in the sequences occur, the distribution of  $k_s$  and  $k_a$  values was calculated (Fig. 5. b, c). For the synonymous mutations, no significant differences occur between the *fmoA*, *bchg* and *sigA* gene sequences. Thus, at sites where the mutation of a nucleotide does not change the coded amino-acid, the three genes evolved equally. For mutations at non-synonymous sites, the sequences of *fmoA* differed from the other two marker genes by showing more sequence differences at sites where point-mutations lead to a change of the respective amino-acid. The *bchg* and *sigA* gene sequences also differ significantly from each other in their  $k_a$  values, the differences however are only little.

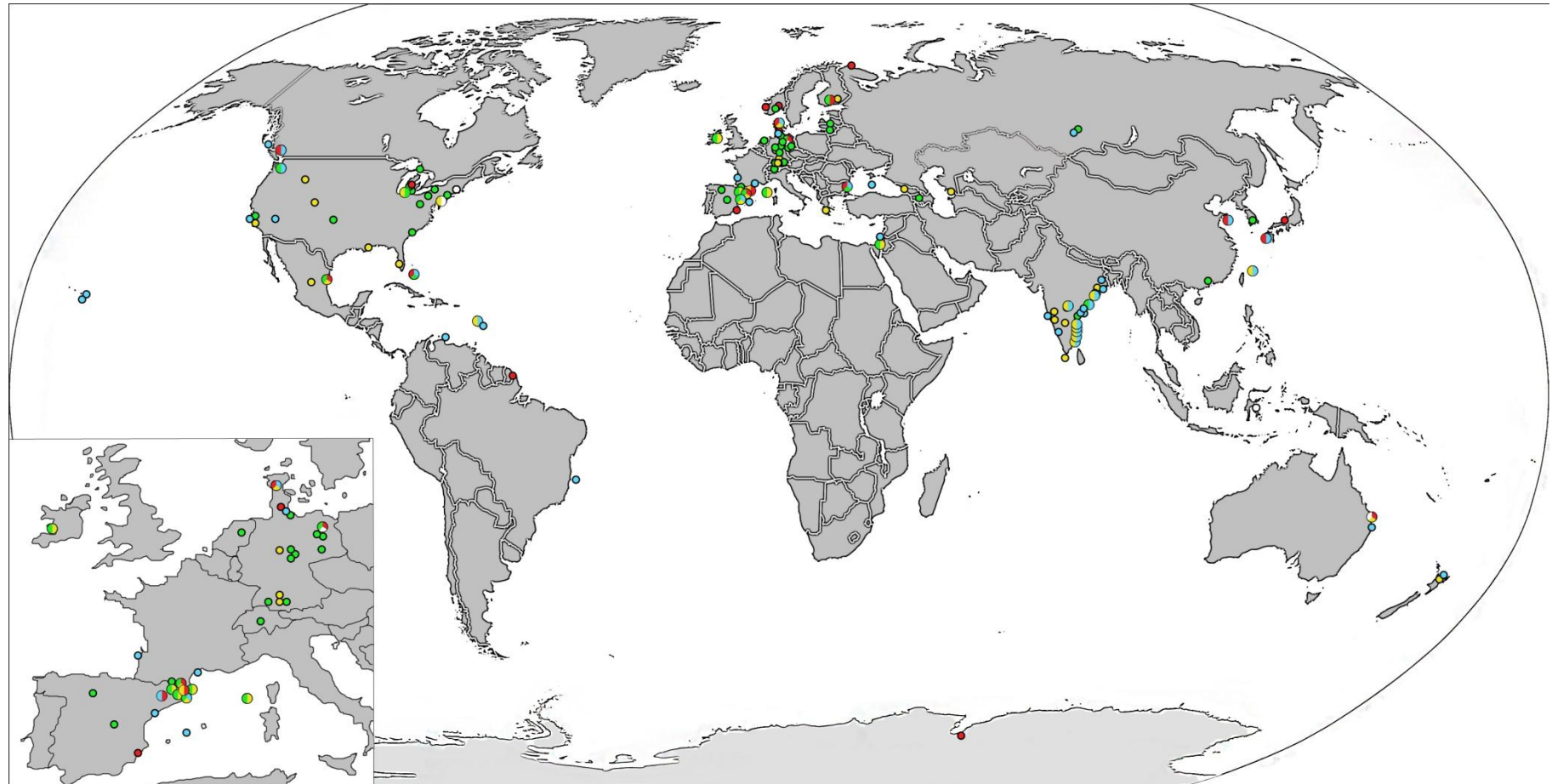




**Figure 5.** The distribution of all pair-wise comparisons of the *fmoA* (—), *bchG* (—) and *sigA* (—) gene sequences plotted as density curves. (a) pairwise dissimilarity values as based on the entire codon, on only the first and second codon position, or only the third codon position. (b) (c) Values of  $k_s$  and  $k_a$ , respectively. The values of the density function sum up to 1, therefore the scaling of the y-axis depends on the values of the x-axis.

### *World map of sampling sites*

The world map containing the sampling sites where green sulfur bacteria have been found (Fig. 6) shows the predominant sampling in Europe, India and North America.



**Figure 6.** World map containing all sampling sites of green sulfur bacteria, color-coded after the four phylogenetic groups suggested by Imhoff (2003). Each sampling site was marked with a dot, which color represents the taxonomic group of green sulphur bacteria (1-4). Split-colored dots represent sampling sites where different groups of GSB were found. Group 1: blue. Group 2: red. Group 3: green. Group 4: yellow. Not assigned: white. The underlying map was obtained from [www.free-vector-maps.com](http://www.free-vector-maps.com).

## 4.2 Results for the investigation of the molecular basis of the interspecies interaction in “*C. aggregatum*”

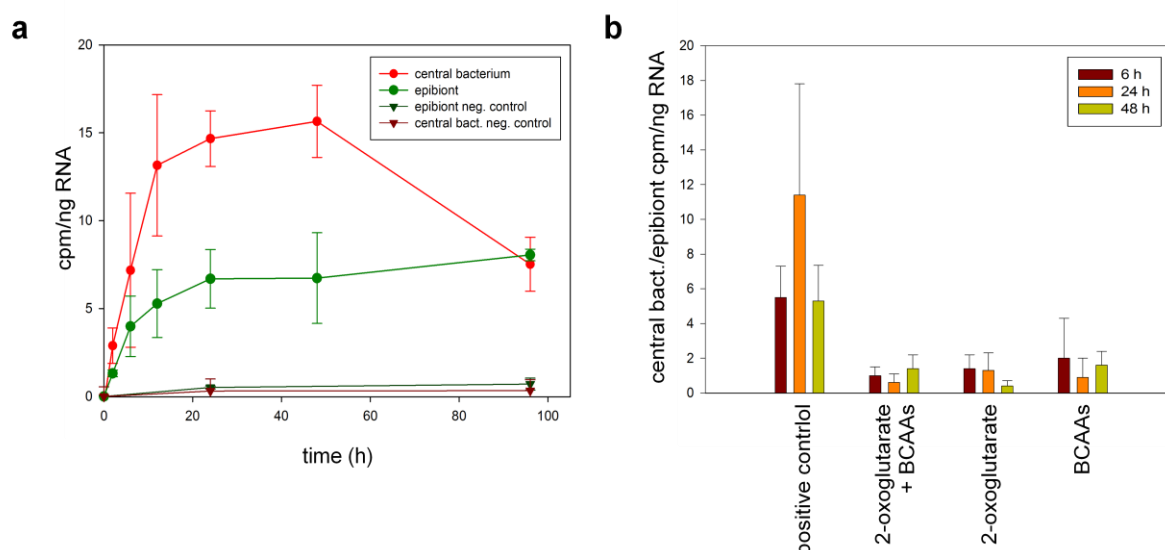
### *Transfer of $^{14}\text{C}$ with magnetic bead capture*

To separate the 16S rRNAs of the epibiont and the central bacterium, we used magnetic beads labeled with specific probes. To verify the success of the separation, cDNA obtained from the RNA isolated was sequenced and analyzed by denaturing gradient gel electrophoresis. The sequencing of the 16S rRNA yielded clear sequences for each sample. A BLAST search (<http://blast.ncbi.nlm.nih.gov>) showed a 100 % match of the sequences with either *Chlorobium chlorochromatii* or the 16S sequence of the central bacterium. On the DGGE gel, only one band per lane was visible (Fig. 7).



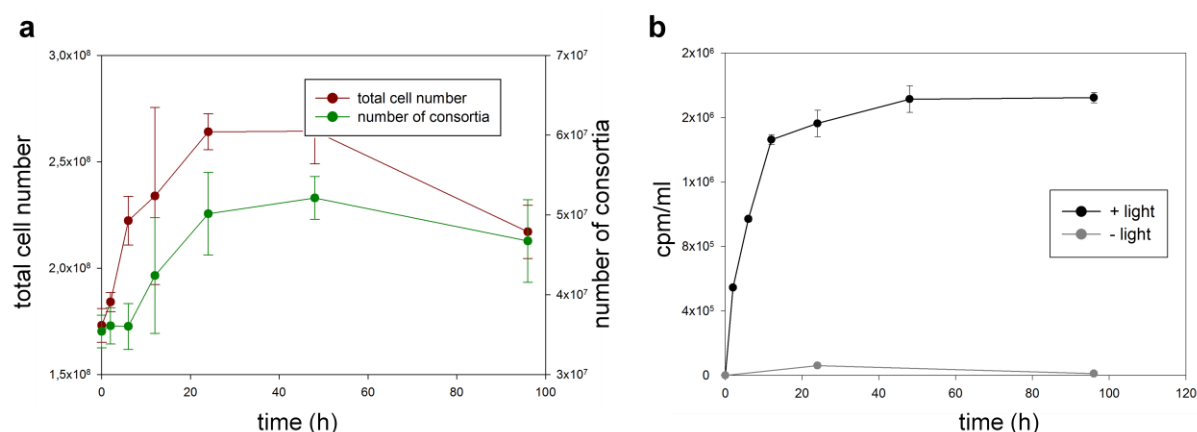
**Figure 7.** Amplification products of enriched epibiont and central bacterium reverse transcribed rRNA on a 30-70 % denaturing gradient gel electrophoresis. eb = epibiont; cb = central bacterium.

From this survey of the magnetic-bead separation-efficiency, we conclude that the RNA captured by magnetic bead separation is of high purity. Therefore, we used this technique to track the incorporation of  $\text{H}^{14}\text{CO}_3^-$  into the partner organisms of “*C. aggregatum*” over a period of 96 h (Fig. 8. a).



**Figure 8.** (a)  $^{14}\text{C}$ -labeling of epibiont and central bacterium rRNA after incubation of the consortia enrichment culture with  $\text{NaH}^{14}\text{CO}_3$  in the light. Negative controls were incubated in the dark. Separation of the specific rRNA was performed by magnetic bead hybridization. (b) Relation between the cpm/ng RNA of the central bacterium and the epibiont after incubation of the consortia enrichment culture with  $\text{NaH}^{14}\text{CO}_3$  in the light for 6 h, 24 h and 48 h. Separation of the specific rRNA was performed by magnetic bead hybridization.

The radioactive label increased simultaneously in both partners, with higher activity in the central bacterium compared with the epibiont. After 48 h, the counts per minute (cpm) / ng RNA dropped for the central bacterium. Hardly any label was measured in the negative controls incubated in the dark. We therefore conclude that a rapid exchange of newly synthesized small molecular weight organic matter occurs from the epibiont to the central bacterium. Further subsampling to increase the resolution during the first two hours of the experiment has not been possible because the values would have fallen below the resolution limit.

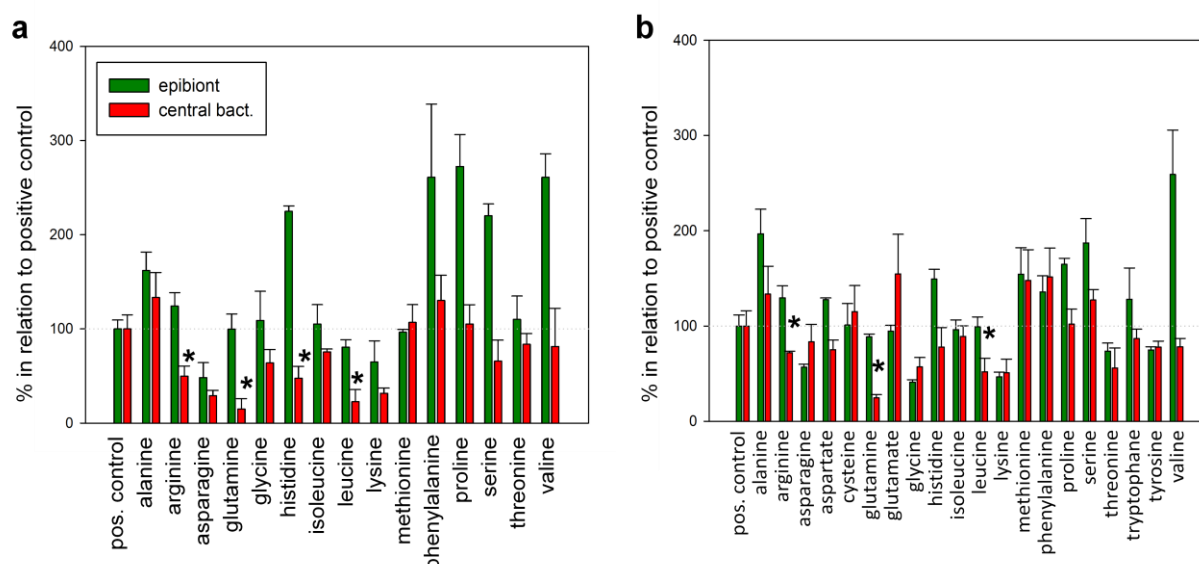


**Figure 9.** (a) Total cell number and number of consortia during incubation of the consortia enrichment culture in the light with  $\text{NaH}^{14}\text{CO}_3$ . Total cell numbers were counted after DAPI, consortia after FISH staining. (b)  $^{14}\text{C}$ -activity of a total cell extract of a consortia enrichment culture after incubation with  $\text{NaH}^{14}\text{CO}_3$  in the light and in the dark.

Total cell numbers increased exponentially, reaching the stationary phase after 24 h with  $(264 \pm 8.46) \cdot 10^6$  cells/ml and decreasing after 48 h (Fig. 9. a). Phototrophic consortia grew after a lag phase of 6 h, peaking at 48 h with  $5.21 \cdot 10^7 \pm 2.68 \cdot 10^6$  consortia/ml before dropping in numbers to  $(46.7 \pm 5.18) \cdot 10^6$  consortia/ml at 96 h. No significant growth was detected in the negative controls. The radioactive label measured in the cell extracts also increased exponentially and reached a plateau at 48 h of incubation with  $(171 \cdot 10^6 \pm 8.21) \cdot 10^4$  cpm/ml. (Fig. 9. b). To better understand the nature of the carbon exchange, unlabeled substrates were used to investigate their effect on the  $\text{H}^{14}\text{CO}_3^-$  incorporation. The addition of 2-oxoglutarate, branched-chain amino acids and a combination of both, each resulted in a strong inhibition of radiolabel incorporation into the central bacterium compared with the positive control (Fig. 8. b).

To identify further substrates able to inhibit the transfer between the two partners, all 20 proteogenic amino acids were separately used as unlabeled substrates (Fig. 10. a). Of those, four were able to decrease the radiolabel in the central bacterium compared with the positive control without decreasing the incorporation into the epibiont as well (Fig. 10. a, bars marked with “\*”). The strongest decrease was measured in the presence of glutamine ( $14.8 \pm 11.1$  %). A complete inhibition of growth for both the epibiont and the central bacterium was caused by the amino acids aspartate, cysteine, glutamate, tryptophane and tyrosine (data not shown). A slightly inhibiting effect has been measured for leucine, lysine and asparagine. On the other

hand, the amino acids alanine, histidine, phenylalanine, proline, serine and valine were able to increase the uptake of radiolabel into the epibiont up to  $272.5 \pm 33.9$  % in the case of proline. Except for alanine, the incorporation of  $^{14}\text{C}$  into the central bacterium in those samples is reduced in comparison with the epibiont. Since this effect could have been caused by contaminating bacteria in the enrichment culture cycling sulfate to sulfide, thus enhancing the epibiont growth, molybdate was added to inhibit sulfate reduction (Yadav and Archer 1989).

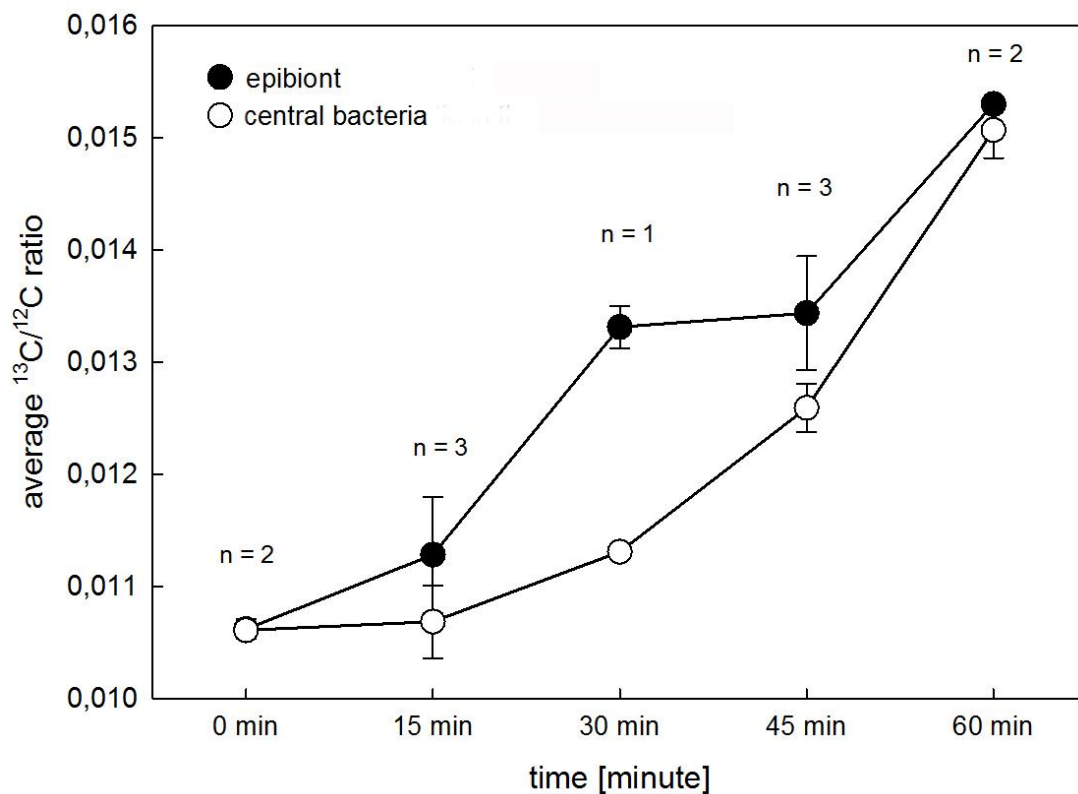


**Figure 10.** Relative  $^{14}\text{C}$ -labeling of epibiont and central bacterium rRNA in relation to the positive control. Consortia were incubated with light,  $\text{NaH}^{14}\text{CO}_3$  and (a) 1 mM amino acid concentration. (b) 0.3 mM amino acid concentration, 1 mM molybdate. (\*)  $^{14}\text{C}$ -label of central bacterium rRNA significantly ( $p < 0.05$ ) reduced in comparison with positive control while epibiont label is comparable to the positive control.

To avoid the growth inhibition that might have been caused by an overconcentration of amino acids, the experiment was repeated with an amino acid concentration of 0.3 mM. This had the desired effect on the growth of phototrophic consortia (Fig 10. b). Although the measured radiolabel in histidine, phenylalanine and proline decreased, the uptake of radioactive  $\text{H}^{14}\text{CO}_3^-$  was still significantly increased in several samples, with the highest value for valine ( $259 \pm 46.6$  %) so that cycling of sulfide very likely did not play a role in the increased radiolabel uptake. The inhibition of carbon transfer could again be shown for arginine, leucine and glutamine with the latter having the strongest effect ( $24.6 \pm 3.4$  %).

### NanoSIMS analysis

Tracking of the  $^{13}\text{CO}_2$  by NanoSIMS revealed an incorporation of the label into both partner organisms. The label could already be detected in the epibiont after 15 min (Fig. 11.). In the central bacterium however, a significant increase in  $^{13}\text{C}$  was only detected after 30 min of incubation. This fast exchange of carbon confirms the conclusion of the magnetic capture experiment that small molecular weight organic matter is transferred between the symbiotic partners. As expected from the  $^{14}\text{C}$ -experiments, the concentration of the label increased over time.

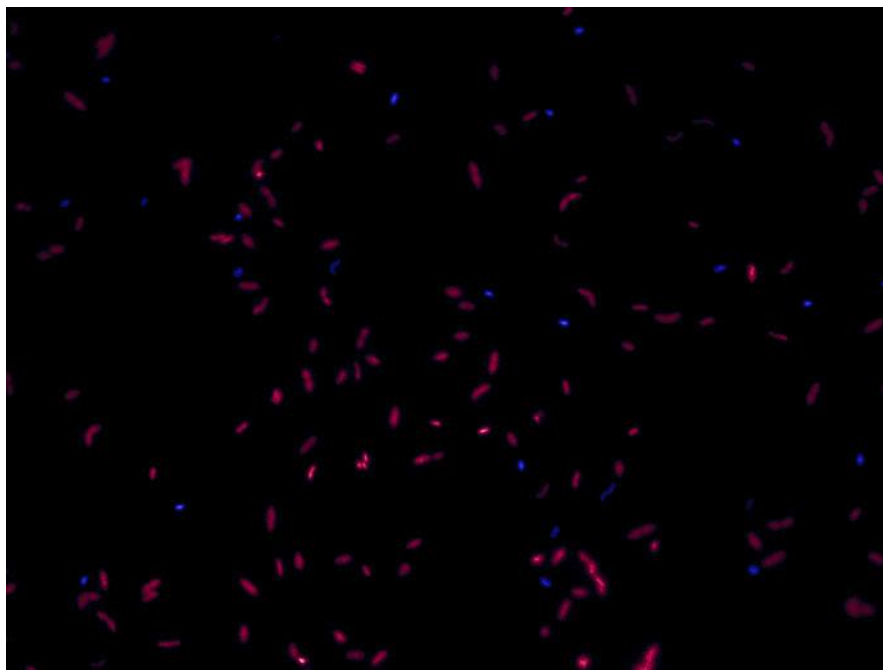


**Figure 11.** Analysis of the  $^{12}\text{C}/^{13}\text{C}$  relation revealing the incorporation of  $\text{H}^{13}\text{CO}_3^-$  into the central bacterium and epibiont by nano-scale secondary ion mass-spectrometry. n = number of consortia analyzed at each time point.

### Isotopologue analysis

To investigate a possible amino acid transfer between the epibiont and the central bacterium,  $^{13}\text{C}$  labeling experiments were conducted. Therefore, the two partners were separated by CsCl density centrifugation (Fig. 12) and the amino acids were analyzed by GC/MS. Results in which labeling of the amino acids with  $^{13}\text{C}$  is below 0.5 % are hardly distinguishable from the

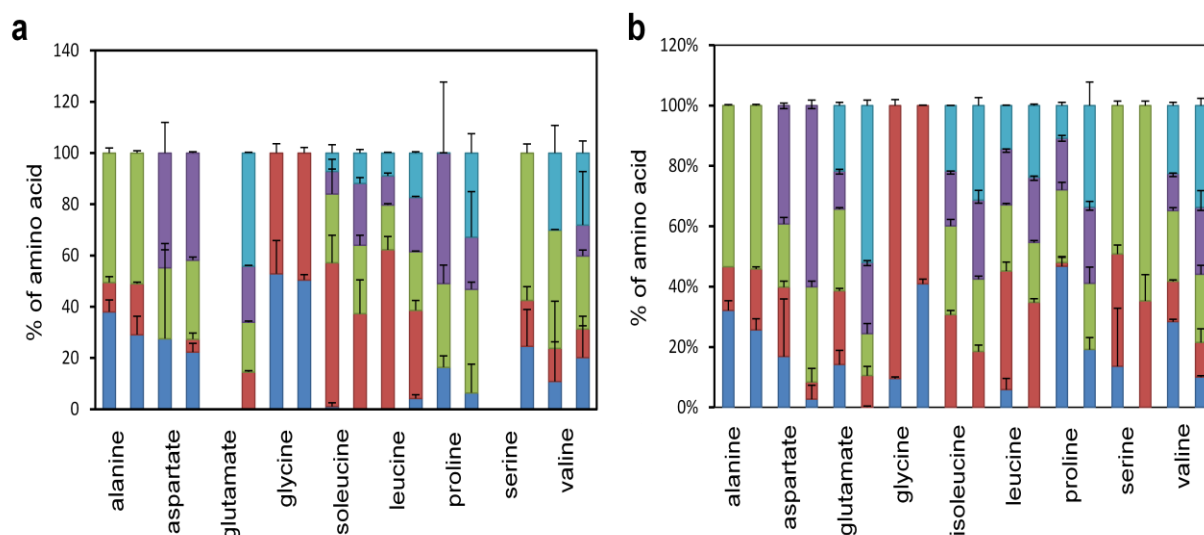
background signal and are therefore not taken into account for the discussion. The analysis of phenylalanine and threonine provided results with very high standard deviations for which reason they are not shown in the results.



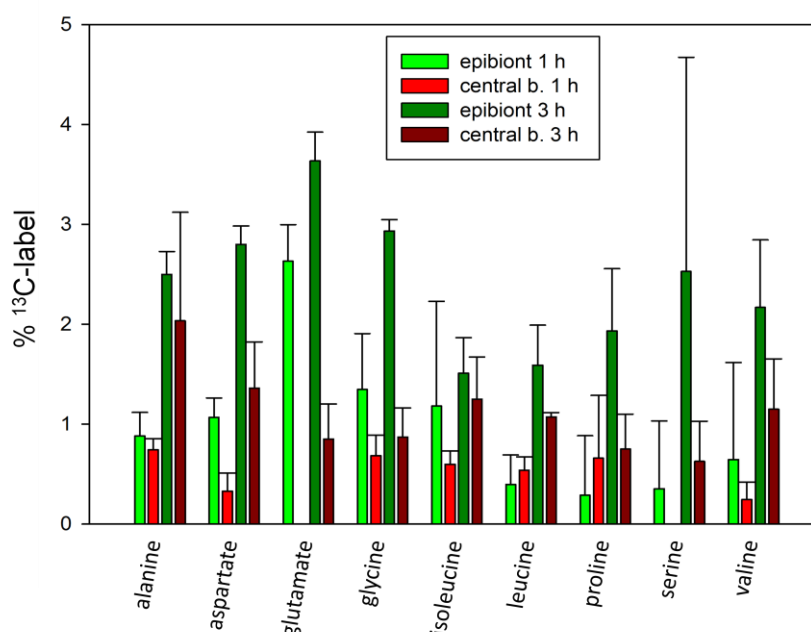
**Figure 12.** Enriched central bacteria after Cs-Cl density gradient centrifugation. The picture is an overlay of DAPI staining (blue) and Cy3 staining (red) by fluorescence *in-situ* hybridization using central rod specific probes (ZS442).

The incubation of consortia with  $\text{H}^{13}\text{CO}_3^-$  resulted in similarly labeled amino acids in the central bacterium and the epibiont after 1 h of incubation (Fig. 13. a). Labeled glutamate and serine were not present in the central bacterium after 1 h of incubation and could therefore not be analyzed. Slight differences occurred in the composition of leucine and proline. Deviations in the labeling pattern increased after 3 hours of incubation (Fig. 13. b). However, the labeling patterns were still comparable to each other. We therefore conclude that the central bacterium is receiving amino acids from the epibiont. The increasing variation between the two labeling patterns over time suggests that a biosynthesis of amino acids is occurring in the central bacterium as well, possibly with other substrates derived from the epibiont.





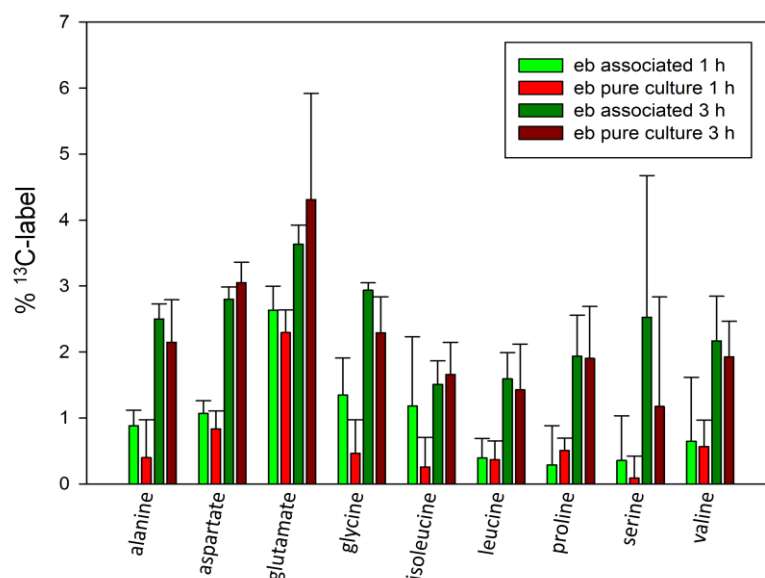
**Figure 13.**  $^{13}\text{C}$ -labeling pattern of central bacterium (left column) and epibiont (right column) amino acids after incubating the consortia enrichment culture for (a) 1 h and (b) 3 h with  $\text{NaH}^{14}\text{CO}_3$  in the light. Cells were separated by  $\text{CsCl}$  density gradient centrifugation and analyzed by GC/MS. Labeled glutamate and serine was not present in the central bacterium and therefore not analyzed. Number of labeled carbon atoms are displayed from bottom to top ( $M+1$ - $M+5$ ).



**Figure 14.** Percentage of  $^{13}\text{C}$ -label in amino acids in the central bacterium and the epibiont after incubating the consortia enrichment culture for 1 h and 3 h with  $\text{NaH}^{13}\text{CO}_3$  in the light. Cells were separated by  $\text{CsCl}$  density gradient centrifugation and analyzed by GC/MS.

In contrast to the labeling pattern, the quantity of the labeled amino acids in the two partner organisms differed (Fig. 14). Glutamate i.e., which was highest labeled in the epibiont after 1 h ( $2.63 \pm 0.37$  %) and 3 h ( $3.63 \pm 0.29$  %), was not present in the central bacterium after 1 h, but only after 3 h ( $0.85 \pm 0.35$  %) of incubation. Thus, the quantity of amino acids synthesized in the epibiont does not necessarily influence the quantity of amino acids transferred to the central bacterium. To investigate the uptake and incorporation of substrates into the central bacterium under the standard cultivation conditions used for phototrophic consortia with 10 mM  $\text{NaHCO}_3$  as opposed to the cultivation with 30  $\mu\text{M}$   $\text{NaHCO}_3$  used in the  $^{14}\text{C}$ -labeling experiments,  $[\text{U-}^{13}\text{C}]$  2-oxoglutarate and  $[\text{U-}^{13}\text{C}]$  leucine was added to the consortia culture. Incorporation of  $^{13}\text{C}$  was only measured in a fraction of the amino acids after the incubation with light and  $\text{NaHCO}_3$ . In the assay in which  $[\text{U-}^{13}\text{C}]$  leucine was added, a high  $^{13}\text{C}$ -label was measured in the amino acid leucine ( $1.43 \pm 0.03$  % in the central bacterium and  $1.75 \pm 0.03$  % in the epibiont), which was almost completely marked with  $^{13}\text{C}$ . Thus, the labeled leucine was taken up and incorporated into the proteins by both organisms. In the assays in which  $[\text{U-}^{13}\text{C}]$  2-oxoglutarate was added, similarly high values as for leucine were measured for isoleucine. Of those,  $38.6 \pm 23.1$  % of the amino acids were completely labeled in the central bacterium and  $28.0 \pm 0.96$  % in the epibiont. The high concentration of  $^{13}\text{C}$ -label and the high amount of completely labeled isoleucine in comparison to the other amino acids suggest that 2-oxoglutarate was partly used to synthesize the labeled isoleucine. Interestingly, glutamate is  $^{13}\text{C}$ -labeled in the epibiont in the assay using  $[\text{U-}^{13}\text{C}]$  2-oxoglutarate as well. Since the labeling pattern shows that glutamate is not completely marked with  $^{13}\text{C}$ , fractions of the  $^{13}\text{C}$ -label should also be found in the other amino acids, but are below the detection limit.

In the experiment conducted in the dark, no incorporation of  $^{13}\text{C}$ -label into the amino acids of the central bacterium or the epibiont could be observed (data not shown).

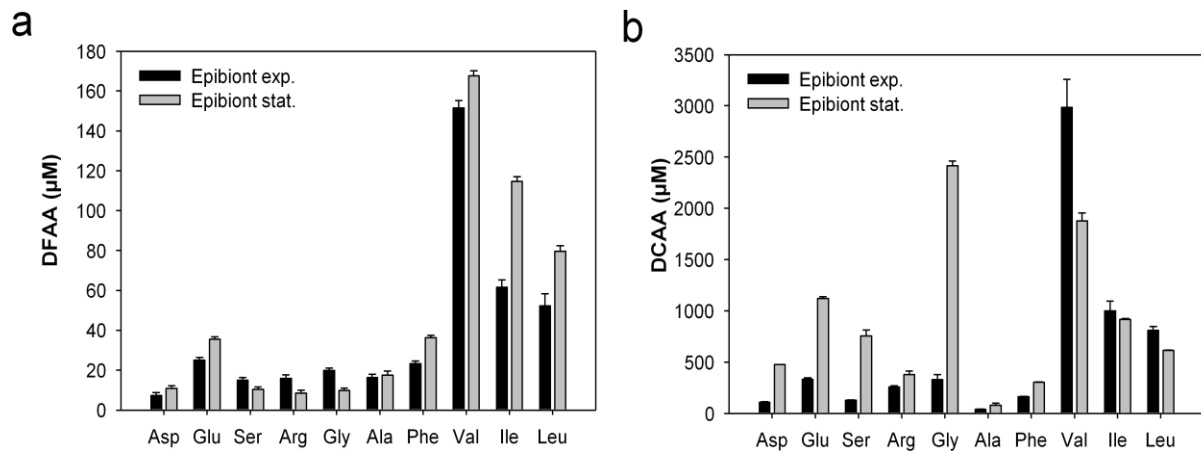


**Figure 15.** Percentage of  $^{13}\text{C}$ -label in amino acids in the epibiont (eb) pure culture and the consortia enrichment culture after incubation for 1 h and 3 h with  $\text{NaH}^{13}\text{CO}_3$  in the light. Cells were separated by  $\text{CsCl}$  density gradient centrifugation and analyzed by GC/MS.

The analysis of the epibiont pure culture amino acids after incubation with  $\text{H}^{13}\text{CO}_3^-$  showed no significant differences to the amino acids of the epibiont in the associated state (Fig. 15). The labeling patterns (data not shown) and the quantity of labeled amino acids were nearly identical. A symbiosis specific adaptation of amino acid production could thus not be observed.

#### *Excretion of amino acids by epibiont cells*

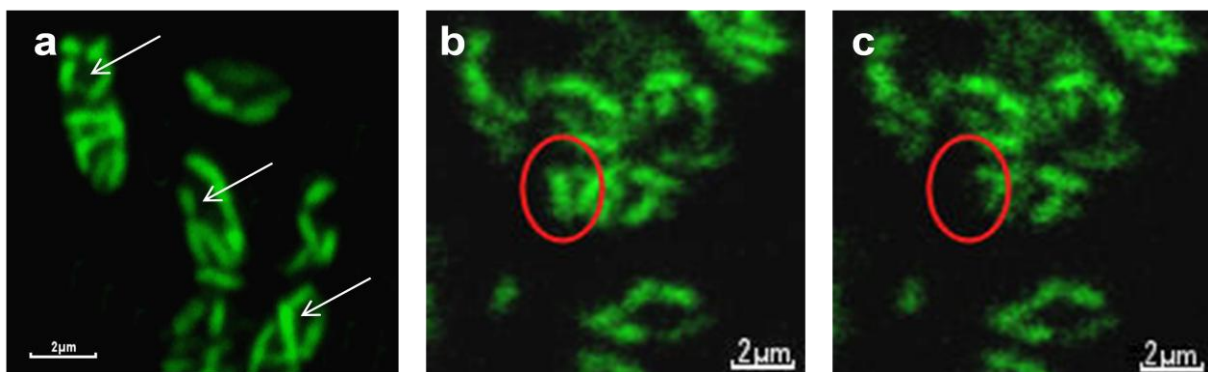
HPLC analysis of dissolved free amino acids (DFAA) revealed that the branched-chain amino acids leucine, valine and isoleucine dominated in supernatants of exponential and stationary epibiont pure cultures (Fig. 16. a). These amino acids thus are excreted by epibiont cells independent of its growth phase. DFAA concentrations of branched-chain amino acids of the stationary phase supernatant were 371  $\mu\text{M}$  valine, 79  $\mu\text{M}$  leucine and 114  $\mu\text{M}$  isoleucine. Parallel analyses of peptides (dissolved combined amino acids (DCAA)) in the culture supernatants revealed that valine, isoleucine and leucine dominated in the DCAA excreted during the exponential growth phase whereas glycine and valine dominated in DCAA of stationary phase epibiont cultures (Fig. 16. b). Overall, concentrations of DCAA were higher by at least one order of magnitude.



**Figure 16.** Amino acid excretion of an epibiont pure culture in the exponential and stationary phase. Concentration of (a) dissolved free amino acids (DFAA) and (b) dissolved combined amino acids (DCAA) were determined by HPLC analysis.

#### *Fluorescence-Recovery-After-Photobleaching (FRAP) analysis*

The imaging of the stained phototrophic consortia revealed that only the epibiont but not the central bacterium had taken up and processed the calcein-AM dye (Fig. 17. a). Therefore,



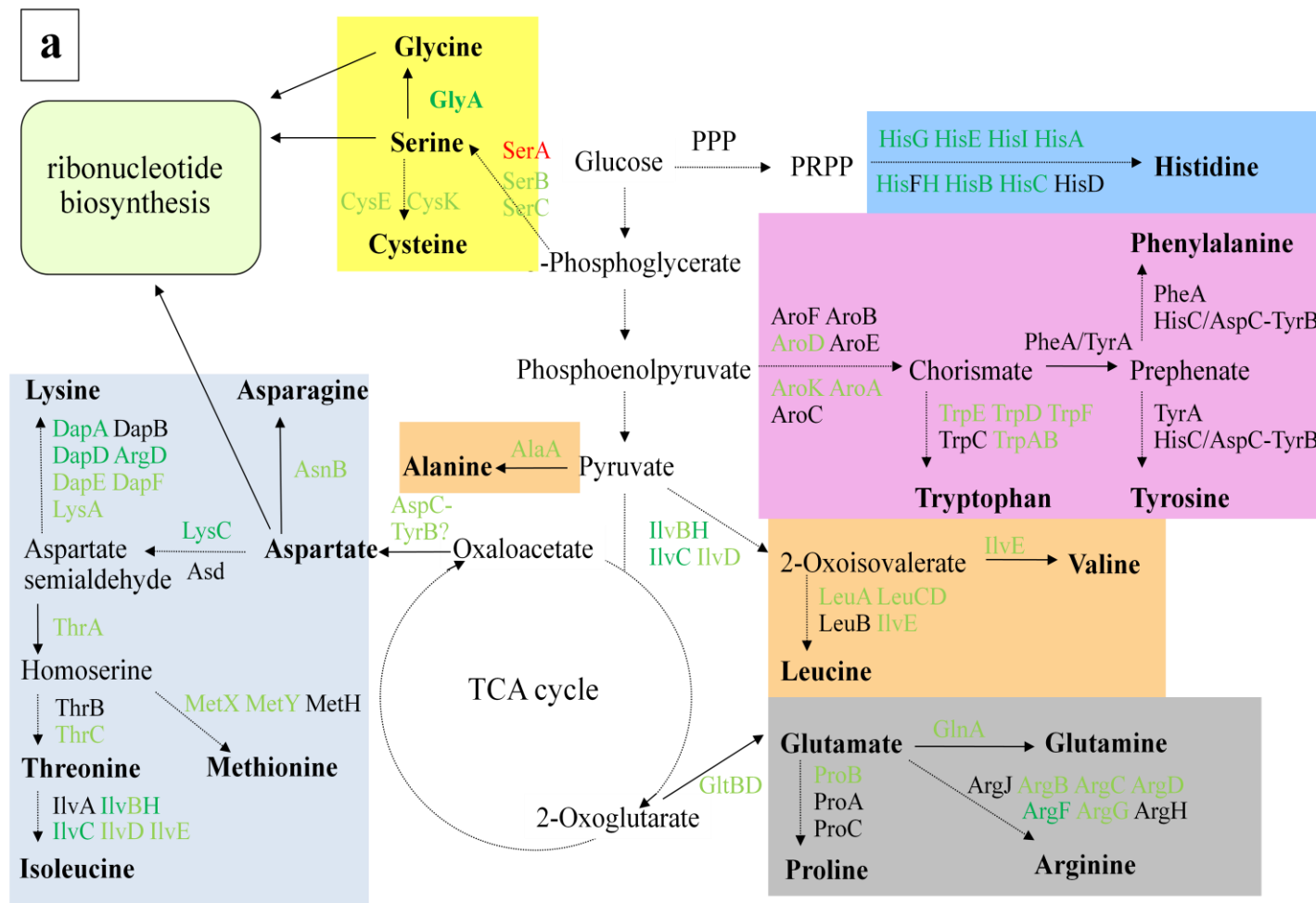
**Figure 17.** “*C. aggregatum*” stained with calcein-AM imaged with a confocal laser scanning microscope at a wavelength of 488 nm. Arrows point towards non-stained central bacteria (a). The red circle marks the epibiont cell being bleached before (b) and 30 seconds after (c) bleaching.

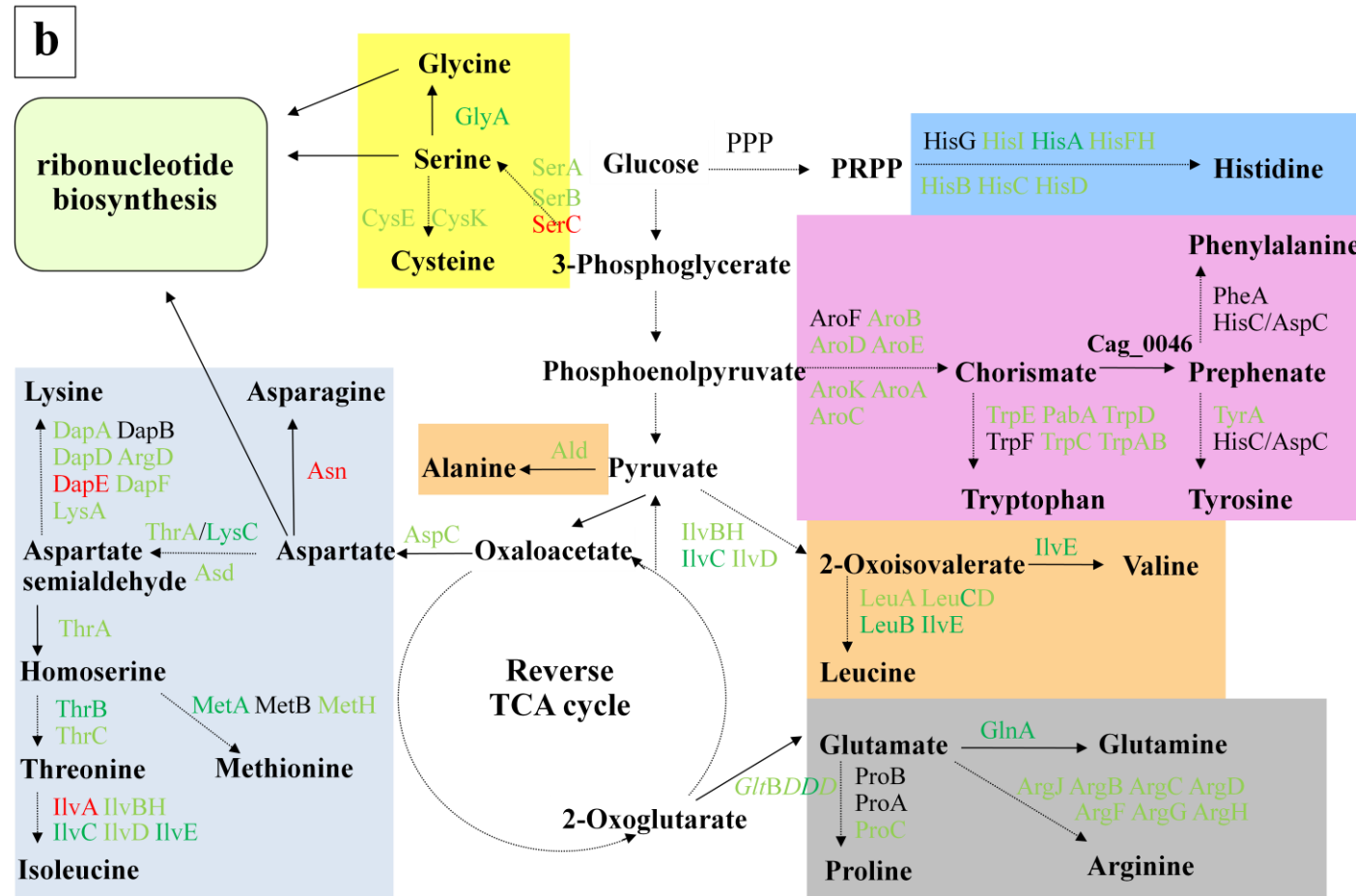
the possibility of free diffusion between the two partner organisms could already be ruled out. Nevertheless, an epibiont was bleached to see if the interconnecting pili between the epibionts might promote diffusion. No fluorescence recovery was observed however after bleaching (Fig. 17. b, c)

### ***Genome analysis***

The complete genome of the central heterotrophic partner bacterium of the phototrophic consortium "*Chlorochromatium aggregatum*" was determined to be a single circular chromosome of 2,991,845 bp including 2772 open reading frames. It bears the characteristics of an aero- or microaerophilic heterotrophic bacterium. It possesses a complete respiratory chain with a NADH:ubiquinone oxidoreductase, a cytochrome *bd* quinol oxidase, a succinate dehydrogenase and a FOF1-type ATP synthase. Furthermore, the genes necessary for the citrate cycle and the pentose phosphate cycle are present. All genes of the glycolysis are encoded in the genome except for the first enzyme phosphorylating  $\beta$ -D-glucose, the glucokinase. No pathways for anoxygenic respiration are encoded in the central bacterial genes. Likewise, of the fermentative pathways known to date, none can be reconstructed using the open reading frames (ORFs) found in the central bacterial genome. Only the possibility to form acetate from pyruvate and CO<sub>2</sub> via pyruvate-ferredoxin oxidoreductase, phosphotransacetylase and acetate kinase exists for the central bacterium. A hydrogenase is present to regenerate the reduced ferredoxin and release molecular hydrogen.

The central bacterial genome encodes all genes required to synthesize the 20 proteinogenic amino acids with the exception of the D-3-phosphoglycerate dehydrogenase *serA* and the aspartate aminotransferase *tyrB* (Fig. 18. a). In addition, a multitude of transporters are found in the central bacterium (Table 6). Of special interest are an ABC-type branched-chain amino acid transport system (CEROD0106-0109, CEROD2035) and two ABC-type oligopeptide transport systems (CEROD0133-0137 and CEROD0986-0989). An ABC-type oligopeptide and a branched-chain amino acid transporter are also encoded in the *Chl. chlorochromatii* genome (Table 7).





**Figure 18.** Amino acid biosynthesis pathways of (a) the central bacterium and (b) the associated epibiont linked to the tricarboxylic or reverse tricarboxylic acid cycle and the ribonucleotide biosynthesis. The genes of the pathways are color-coded according to their RPKM-values of the transcriptome analysis with light green indicating low transcription (RPKM between 1-10), dark green indicating high transcription (RPKM > 10), black indicating no or nearly no transcription (RPKM < 1) and red marking genes not found in the genome. Dashed lines mark pathways in which more than one enzyme is engaged.

**Table 6.** RPKM-values of the transporters of the central bacterium incubated in the light and in the dark with the addition of BCAAs and 2-oxoglutarate.

ABC-type transporter			
annotation	gene name	RPKM + light	RPKM - light
CEROD0679	sugar transport system, periplasmic component/surface lipoprotein	1.50	0.00
CEROD0555	sugar transport system, periplasmic component	0.33	0.00
CEROD1125	ribose transport system, periplasmic component	12.26	0.37
CEROD0642	ribose transport system, substrate-binding protein	1.66	0.18
CEROD0578	lipoprotein-releasing system	0.35	0.00
CEROD0579		1.81	0.00
CEROD0133	oligopeptide transport system	183.32	7.63
CEROD0134		51.66	8.38
CEROD0135		70.19	5.43
CEROD0136		6.67	0.56
CEROD0137		4.68	0.44
CEROD0986	oligopeptide transport system	0.00	0.00
CEROD0987		0.00	0.00
CEROD0988		0.84	0.00
CEROD0989		0.48	0.00
CEROD2496	antimicrobial peptide transport system	0.00	0.00
CEROD2240	molybdate transport system, substrate-binding protein	49.63	0.23
CEROD2509	cobalt/nickel transport system, permease protein	4.86	0.17
CEROD0588	multidrug transport system, ATPase component	0.60	0.00
CEROD0589		1.72	0.24
CEROD0551	cation/multidrug efflux pump	2.36	0.45
CEROD0671	phosphate transport system <i>pstSCAB</i>	0.00	0.00
CEROD0672		0.00	0.00
CEROD0673		0.42	0.00
CEROD0674		1.92	1.96
CEROD2476	phosphonate transport system, substrate-binding protein	6.41	0.84
CEROD0180	lipid carrier protein	0.00	0.12
CEROD0106	branched-chain amino acid transport system	0.00	0.00
CEROD0107		0.00	0.00
CEROD0108		2.00	1.01
CEROD0109		0.46	0.39
CEROD2027	branched-chain amino acid transport system,	1.43	0.15



	periplasmic component		
CEROD2059	branched-chain amino acid transport system, periplasmic component	5.49	6.98
CEROD2183	branched-chain amino acid transport system, ATPase component	0.60	0.00
CEROD2266	branched-chain amino acid transport system, permease component	0.40	0.00
CEROD2267		2.34	0.74
CEROD2036	amino acid transport/signal transduction system, periplasmic component/domain	0.60	0.00
CEROD0184	cobalamin/Fe3 <sup>+</sup> -siderophores transport system	0.00	0.00
CEROD0185		0.00	1.98
CEROD0186		8.03	0.49
CEROD1894	Fe3 <sup>+</sup> -hydroxamate transport system	0.58	0.24
CEROD1895		0.39	0.00
CEROD1896		0.44	0.00
CEROD0041	transport system involved in resistance to organic solvents	25.04	0.41
CEROD0042		45.51	1.12
CEROD0043		20.56	7.53
CEROD0044		18.06	1.61
CEROD0794	ABC-type transport system, ATPase component	0.53	0.00
CEROD1737	ABC-type uncharacterized transport system, permease and ATPase component	0.71	0.00
CEROD2284	protease/lipase transport system, ATPase and permease components	80.53	0.00
CEROD0688	protease/lipase transport system, ATPase and permease components	2.82	0.30
CEROD1472	sulfonate/nitrate/taurine transport system, substrate-binding protein	29.57	2.33
CEROD2290	polysaccharide/polyol phosphate transport system, ATPase component	7.66	0.00
CEROD2291		33.80	0.00
CEROD2193	putrescine transport system, substrate-binding protein	0.51	0.00
CEROD2194		5.63	0.20
CEROD2195		0.79	0.00
CEROD2196		11.03	0.75
CEROD1936	heme exporter	0.00	0.00
CEROD1937		0.00	0.00
CEROD1938		0.00	0.00
CEROD0560	ABC-type transport system of unknown function	2.95	0.47
CEROD0561		6.72	0.00
non-ABC-type ATPase			
CEROD1756	F-type H <sup>+</sup> -transporting ATPase	28.80	2.31

CEROD1757		114.89	2.21
CEROD1758		40.49	1.17
CEROD1759		8.16	1.04
CEROD1760		6.97	0.24
CEROD1751		0.93	0.20
CEROD1752		2.73	0.26
CEROD1763		1.04	0.00
CEROD1741	ATPase of AAA+ superfamily	0.36	0.00
<b>ion channels</b>			
CEROD0621	magnesium and cobalt transporter	0.50	0.21
CEROD1878	large conductance mechanosensitive channel	5.74	3.24
<b>secondary transporter</b>			
CEROD2307	arsenite efflux pump ACR3 and related permeases	1.83	1.94
CEROD1929	Na <sup>+</sup> /H <sup>+</sup> antiporter	2.01	0.00
CEROD1853	preprotein translocase	19.08	0.38
CEROD1854		3.01	0.10
CEROD1855		39.67	2.25
CEROD0744	sec-independent protein translocase	7.71	0.82
CEROD0745		0.00	0.00
CEROD0746		0.00	0.00
<b>unclassified</b>			
CEROD2035	biopolymer transport protein	8.06	1.42
CEROD2198	biopolymer transport protein	0.00	0.00
CEROD1386	biopolymer transport protein	0.63	0.00
CEROD1387		0.00	0.00
CEROD1388		0.00	0.00
CEROD2022	lipopolysaccharide export system	0.39	0.00
CEROD2023		0.80	0.00
CEROD2520	long-chain fatty acid transport protein	3.55	0.00
CEROD0533	Co <sup>2+</sup> /Zn <sup>2+</sup> /Cd <sup>2+</sup> cation transporter	0.45	0.19

**Table 7.** RPKM-values of the transporters of the epibiont *Chlorobium chlorochromatii* CaD3<sup>T</sup> in the associated state incubated in the light and in the dark with the addition of BCAAs and 2-oxoglutarate. Annotation of transporters was adopted from TransportDB (<http://www.membranetransport.org>).

annotation	transporter type	RPKM symbiotic + light	RPKM symbiotic - light
<b>ABC-type transporters</b>			
Cag_1498	nucleoside transporter	13.95	1.36
Cag_1499		9.40	1.23
Cag_0156	nucleoside transporter, membrane protein	105.15	61.00
Cag_0792	nucleoside transporter, binding protein	294.36	11.31
Cag_0264	lipoprotein-releasing system, ATP-binding protein	36.41	3.40
Cag_0371	integral membrane protein	36.55	1.95
Cag_0869	macrolide-specific ABC-type efflux carrier, ATPase component	165.62	28.85
Cag_0919	lipoprotein release transport system, permease component	36.38	1.59
Cag_0920		14.50	1.20
Cag_1781	lipoprotein releasing system	38.95	1.42
Cag_0337	macrolide export, ATP-binding/permease protein	34.96	2.76
Cag_0413	molybdate transport system	160.18	6.65
Cag_0414		50.91	1.20
Cag_0415		30.25	2.73
Cag_1705	oligopeptide/dipeptide transporter	91.05	6.22
Cag_1157	oligopeptide/dipeptide transporter	21.54	1.74
Cag_0404	oligopeptide/dipeptide transporter	16.55	3.98
Cag_1883	oligopeptide/dipeptide transporter	34.46	3.39
Cag_1765	oligopeptide/dipeptide transporter	46.43	5.64
Cag_0453	multidrug transport system, ATPase and permease component	37.95	3.64
Cag_1327	multidrug transport system, ATPase and permease component	36.89	8.00
Cag_0498	phosphate import transporter	11.35	0.05
Cag_0499		13.11	0.94
Cag_0500		26.13	1.52
Cag_0501		58.07	2.53
Cag_0849	brached-chain amino acid transporter	60.75	2.31

Cag_0850		54.69	2.60
Cag_0851		225.10	7.62
Cag_0852		460.16	14.98
Cag_0853		1535.66	29.01
Cag_1079	cobalamin/Fe <sup>3+</sup> -siderophores transport system	49.35	9.97
Cag_0754	siderophore transport system, ATP-binding cassette	23.21	2.03
Cag_0726	cobalamin binding protein	15.86	2.12
Cag_1674	Zn <sup>2+</sup> /Mn <sup>2+</sup> transport system	38.22	3.92
Cag_1675		80.36	9.17
Cag_1676		31.92	1.10
Cag_0478	organic solvents resistance transport system	214.80	10.37
Cag_0479		49.93	1.85
Cag_0480		35.50	4.15
Cag_0740	type I protease secretion system, ATPase component	152.87	11.27
Cag_0251	unclassified ABC-type transport system, ATPase component	48.06	3.25
Cag_0334	ATPase components of ABC transporters with duplicated ATPase domains	115.94	5.84
Cag_0375	ATPase components of ABC transporters with duplicated ATPase domains	145.46	49.10
Cag_0716	unclassified ABC-type transport system	4.10	0.80
Cag_0717		8.10	1.22
Cag_1482	lipid A export system, transmembrane region	30.18	0.72
Cag_0988	antimicrobial peptide transport system	27.54	1.59
Cag_0989		39.88	5.60
Cag_0990		88.27	16.45
Cag_0117	arsenite, antimonite and arsenate efflux pump	435.15	3.49
Cag_0217	anion-transporting ATPase	321.16	2.81
Cag_0221	anion-transporting ATPase	111.16	3.87
Cag_1378	anion-transporting ATPase	346.90	10.18
Cag_1412	anion-transporting ATPase	128.47	4.58
Cag_0064	F0F1 ATP-synthase	220.94	15.70
Cag_0065		1910.63	83.40
Cag_0066		523.28	11.09
Cag_0067		143.14	5.96

Cag_0140	F0F1 ATP-synthase	1238.78	14.74
Cag_0141		211.85	7.34
Cag_2014	F0F1 ATP-synthase	631.34	15.80
Cag_2015		12.39	0.00
Cag_0443	membrane-bound proton-translocating pyrophosphatase	158.08	5.61
Cag_1090	membrane-bound proton-translocating pyrophosphatase	40.41	1.29
Cag_0805	secretory pathway, signal recognition	243.57	6.51
Cag_1212	secretory pathway, signal recognition	96.59	7.09
Cag_0959	heavy-metal transport associated domain	20.08	2.94
Cag_1369	heavy-metal transport associated domain	12.11	0.76
Cag_1142	cation transport, ATPase component	37.86	1.37
<b>ion channels</b>			
Cag_1999	ammonium transport	25.35	19.59
Cag_0243	glutamate gated ion channel	48.23	3.62
Cag_1581	Mg <sup>2+</sup> /Co <sup>2+</sup> transporter	17.70	1.76
Cag_0025	large conductance mechanosensitive ion channel	52.80	14.45
Cag_0247	small conductance mechanosensitive ion channel	372.73	32.66
<b>phosphotransferase system</b>			
Cag_1869	histidine-containing phosphocarrier protein	21.81	32.66
Cag_1468	fructose/mannitol phosphotransferase system, IIA subunit	100.38	5.60
Cag_1607	phosphoenolpyruvate-protein phosphotransferase	49.21	2.27
<b>secondary transporter</b>			
Cag_0408	sodium/bile acid symporter	18.09	28.97
Cag_0923	amino acid permease	57.41	4.63
Cag_0950	K <sup>+</sup> -dependent Na <sup>+</sup> /Ca <sup>2+</sup> exchanger-related protein	33.12	3.79
Cag_1143	Co <sup>2+</sup> /Zn <sup>2+</sup> /Cd <sup>2+</sup> cation transporter	64.46	13.54
Cag_1400	chloride ion channel	24.10	0.94
Cag_1205	sodium/proton antiporter	10.13	2.61
Cag_1550	K <sup>+</sup> transport system, membrane component	60.29	11.39
Cag_0283	small multidrug resistance protein	19.46	2.80
Cag_1095	small multidrug resistance protein	30.03	1.26
Cag_1725	potassium uptake protein	153.68	11.36

Cag_0039	drug resistance transporter	76.18	9.74
Cag_0889	drug resistance transporter	30.97	1.38
Cag_0936	drug resistance transporter	107.86	9.65
Cag_2010	drug resistance transporter	4.30	0.26
Cag_0725	drug resistance transporter	10.50	0.36
Cag_0424	drug:proton antiporter	18.75	1.62
Cag_0150	transporter of unknown function (Major Facilitator Superfamily (MFS))	27.19	6.81
Cag_1339	sugar transport protein	1109.92	59.28
Cag_0665	Na <sup>+</sup> -driven multidrug efflux pump	43.42	1.00
Cag_1862	polysaccharide efflux transporter	9.13	0.97
Cag_1493	xanthine/uracil transporter	92.69	26.54
Cag_0026	preprotein translocase	311.37	21.45
Cag_0027		126.17	14.54
Cag_0123	cation/multidrug efflux pump	27.31	4.37
Cag_0971	acriflavin efflux	15.71	1.75
Cag_0195	sodium/solute symporter	55.33	5.51
Cag_1553	sulfate anion transporter	19.06	7.97
Cag_0964	Sec-independent periplasmic protein translocase	305.38	33.27
Cag_1748	twin-arginine translocation protein	691.85	70.32
<b>unclassified</b>			
Cag_1929	ferrous ion transport protein	87.03	31.45
Cag_1801	Mg <sup>2+</sup> /Co <sup>2+</sup> /Ni <sup>2+</sup> -transporter	41.96	4.03

In addition to a multitude of transporters, the central bacterial genome encompasses 32 genes coding for proteins possessing a PAS domain. These domains are ubiquitous in signal transduction proteins and often serve as direct sensors of internal or external stimuli. Their spectrum of stimuli ranges from small molecules, ions, gases to light and redox state (Henry and Crosson 2011). Unfortunately, the signal or ligand of only a small fraction of PAS domains annotated have been identified (Möglich *et al.* 2009). In the central bacterium, only a bacteriophytochrome, a light-regulated signal transduction histidine kinase containing a PAS domain has a predicted signal or ligand binding function. This however is of special interest

with regard to "*C. aggregatum*" since it implies that orientation towards light is conducted by the central bacterium and not by the epibiont as postulated (Overmann and Schubert 2002).

### ***In silico subtractive hybridization***

The *in silico* subtractive hybridization revealed the genome of *Chl. chlorochromatii* to be of average size compared to the eleven other available green sulfur bacterial genomes with 2.57 Mbp (Table 8). Interestingly, although it is engaged in a symbiosis with a heterotrophic bacterium, it only possesses 99 unknown unique ORFs in comparison with the average of 186.8 unique ORFs. However, the %-(G+C)-content of its genome is the smallest among the green sulfur bacteria with 44.3%.

**Table 8.** Comparison of green sulfur bacterial genomes against each other. *In silico* subtractive hybridization was conducted with the Phylogenetic Profiler available at the DOE Joint genome Institute website (<http://img.jgi.doe.gov>).

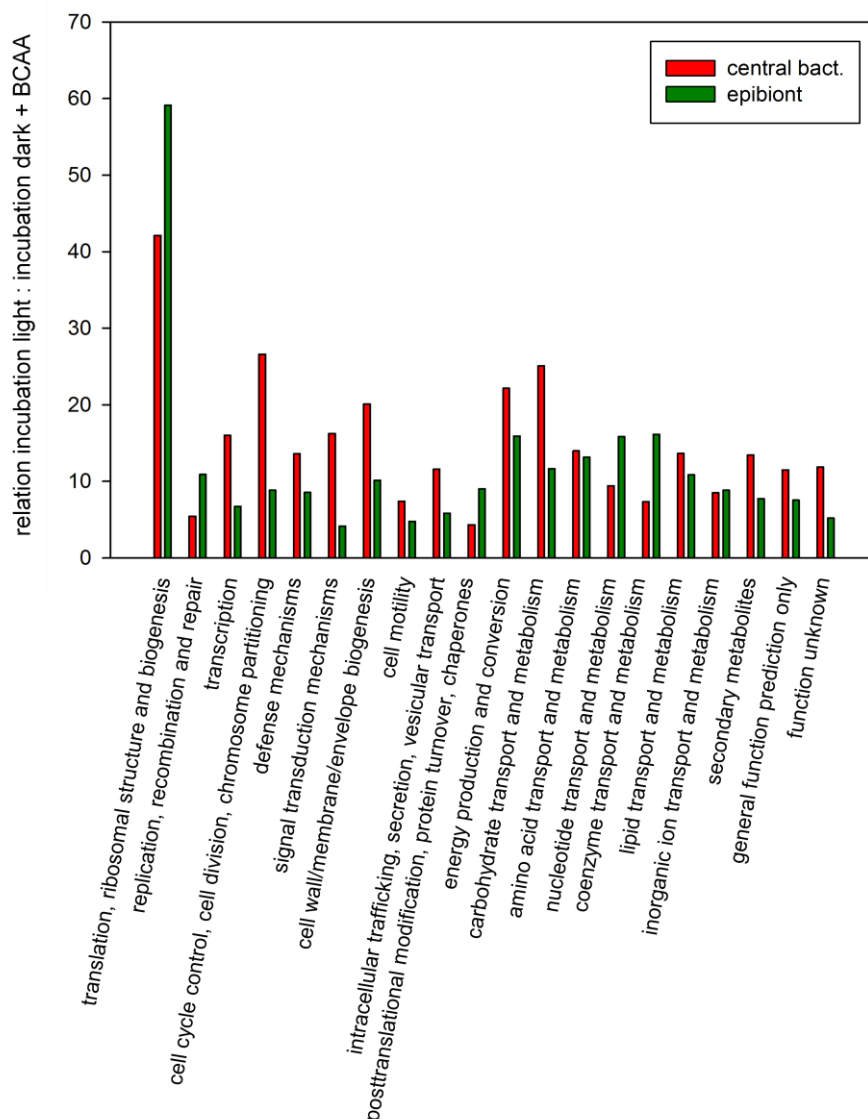
green sulfur bacterium	genome size in Mbp	ORFs-in genome	unique ORFs	%-of genome	unknown unique ORFs	%-(G+C)
<i>Chlorobium chlorochromatii</i> CaD3	2.57	2002	186	9.3	99	44.3
<i>Chlorobaculum parvum</i> NCIB 8327	2.29	2078	139	6.7	81	55.8
<i>Chlorobaculum tepidum</i> TLS	2.15	2252	396	17.6	366	56.5
<i>Chlorobium ferrooxidans</i> DSM 13031	2.54	2158	181	8.4	84	50.1
<i>Chlorobium limicola</i> DSM 245	2.76	2522	204	8.1	149	51.3
<i>Chlorobium phaeobacteroides</i> BS1	2.74	2559	331	12.3	211	48.9
<i>Chlorobium phaeobacteroides</i> DSM 266	3.13	2743	266	9.7	164	48.4
<i>Chlorobium phaeovibrioides</i> DSM 265	1.97	1773	56	3.2	33	53.0
<i>Chloroherpeton thalassium</i> ATCC 35110	3.29	2731	971	35.6	396	45.0
<i>Pelodictyon phaeoclathratiforme</i> BU-1	3.02	2911	550	18.9	394	48.1
<i>Chlorobium luteolum</i> DSM 273	2.36	2083	100	4.8	49	57.3
<i>Prosthecochloris aestuarii</i> SK413, DSM 271	2.58	2402	409	17.0	215	50.1
<b>Average</b>	<b>2.61</b>	<b>2351.2</b>	<b>315.8</b>	<b>12.6</b>	<b>186.8</b>	<b>50.1</b>

### ***Transcriptome analysis***

The transcriptome of consortia incubated under standard conditions in the light and a consortia culture incubated in the dark with the addition of BCAAs and 2-oxoglutarate were analyzed. Overall,  $3.21 \cdot 10^7$  reads were obtained from the consortia culture in the light and  $2.48 \cdot 10^7$  reads from the culture incubated in the dark. Of those reads,  $2.31 \cdot 10^6$  and  $5.45 \cdot 10^5$  reads originated from the central bacterium respectively. For both the central bacterium and the epibiont, the mRNA quantity decreased in the essays incubated in the dark. Based on clusters of orthologous groups of proteins (COGs) (Tatusov *et al.* 2000), the biggest difference in transcription between the incubation of the symbiotic epibiont in the dark and in the light is found in the group "proteins for translation, ribosomal structure and biogenesis", of which the overall RPKM value is increased 59.1 fold (Fig. 19) in the light in comparison with the incubation in the dark. The other groups in comparison are enhanced between 4.2 fold ("signal transduction mechanisms") and 16.1 fold ("coenzyme transport and metabolism").

For the central bacterium, "translation, ribosomal structure and biogenesis" also shows the highest upregulation with a 42.1 fold increase (Fig. 19). The group "posttranslational modification, protein turnover, chaperones" is least increased (4.3 fold). Notably, the RPKM values of each group of the eggNOG classification is increased several fold in the light in comparison with the dark, although BCAAs and 2-oxoglutarate were available. Thus, if the central bacterium should be metabolically active in the dark with carbon substrates in the culture medium, the activity is drastically reduced in comparison with the incubation in the light where the epibiont is able to perform photosynthesis.





**Figure 19.** Increase of RPKM-values of consortia enrichment cultures cultivated in the light in comparison to the dark with the addition of BCAAs and 2-oxoglutarate. The bars depict the relation between the sum of the RPKM-values of the differently incubated cultures based on clusters of orthologous groups of proteins (COGs) (Tatusov *et al.* 2000).

Among the genes that show the highest transcription in the central bacterium, a relatively high amount of uncharacterized and hypothetical proteins is found. Overall, those proteins make up 34.7 % of the genome. But when looking at the 200 genes showing the highest transcription in the mRNA analysis, the proteins with unknown function make up 46.3 % when incubated in the light and 47.1 % when incubated in the dark. The remaining genes are predominantly

involved in energy metabolism, ribosome synthesis and cell division. Unknown symbiosis specific genes might therefore play an important role in the central bacterium.

As has been revealed in the genome analysis, nearly all genes are present to synthesize the 20 proteogenic amino acids in the epibiont and central bacterium respectively. Most of these genes are transcribed at least at a basic level (RPKM value  $>1$ ) (Fig. 18. a, b). The genes missing in the epibiont amino acid biosynthesis pathways are commonly also not present in the 12 green sulfur bacterial genomes available, therefore indicating alternative enzymes taking over the respective part of the synthesis. In the central bacterium, the only two amino acids for which none of the respective genes for biosynthesis are transcribed are tyrosine and phenylalanine. For other amino acids like lysine or histidine, not all genes necessary for synthesis are transcribed. However, the fact that mRNA of the other genes is present suggests that those amino acids are synthesized by the central bacterium itself. Interestingly, similar findings as in the central bacterium can be made in the epibiont (Fig. 18. b). Most genes for the amino acid biosynthesis pathways are transcribed except for phenylalanine and in the case of tyrosine, only one of the two genes necessary for biosynthesis is transcribed. Accordingly, mRNA of only one of the three enzymes of proline synthesis is found in the central bacterium and the epibiont.

The glycolysis and the citrate cycle are both transcribed in the central bacterium (Table 9) incubated in the light. Thus, glycolysis seems to be performed, although the first enzyme of the pathway, the glucokinase is missing.

The complexes of the respiration chain are highly transcribed in the central bacterium (Suppl. Table 9). Thus, also "*C. aggregatum*" is incubated under anoxic conditions, according to transcriptome analysis, the central bacterium seems to possess an active respiratory chain. The high transcription of the respiration chain also explains why the RPKM values of the succinate dehydrogenase, which is the only enzyme active in both the citrate cycle and the respiration chain, is elevated in comparison with the other genes of the citrate cycle.

**Table 9.** RPKM-values of the glycolysis, gluconeogenesis, citrate cycle and respiration chain of the central bacterium incubated in the light and in the dark with the addition of BCAAs and 2-oxoglutarate. The reference organism is *E. coli* K12 and the nearest species with suitable genome is *Rhodospirillum rubrum* T118.

glycolysis and gluconeogenesis			
annotation	gene name	RPKM + light	RPKM - light
-----	glucokinase		
CEROD1141	glucose-6-phosphate isomerase	1.95	0.35
CEROD0898	6-phosphofructokinase	41.00	0.50
-----	fructose-1,6-bisphosphatase		
CEROD1289	fructose-bisphosphate aldolase	2.85	0.17
CEROD1707	triosephosphate isomerase	9.90	0.99
CEROD2151	glyceraldehyde 3-phosphate dehydrogenase	48.87	1.45
CEROD2152	phosphoglycerate kinase	6.17	1.54
CEROD0783	2,3-bisphosphoglycerate-dependent phosphoglycerate mutase	27.96	0.99
CEROD0276	enolase	100.93	1.00
CEROD1021	phosphoenolpyruvate carboxykinase	78.33	3.41
CEROD1022	pyruvate kinase	13.00	2.18
citrate cycle			
annotation	gene name	RPKM + light	RPKM - light
CEROD2158	pyruvate dehydrogenase E1 component	1.76	0.14
CEROD2157	pyruvate dehydrogenase E2 component	5.14	0.15
CEROD2468	citrate synthase	3.31	0.00
CEROD2712	aconitate hydratase 1	0.67	0.21
CEROD1010	isocitrate dehydrogenase	1.73	0.44
CEROD0301	2-oxoglutarate dehydrogenase E1 component	1.90	1.06
CEROD0302	2-oxoglutarate dehydrogenase E2 component	3.07	0.43
CEROD2532	succinyl-CoA synthetase	4.43	1.10
CEROD2531		0	0.00
CEROD2473	succinate dehydrogenase flavoprotein	92.03	3.13
CEROD2472		397.91	0.10
CEROD2471		29.76	9.03
CEROD2470		108.20	2.10
CEROD0338	fumarate hydratase class I	6.34	0.58
CEROD2474	malate dehydrogenase	14.05	0.19

CEROD2468	citrate synthase	3.31	0.00
<b>respiration chain</b>			
<b>succinate dehydrogenase</b>			
annotation	gene name	RPKM + light	RPKM - light
CEROD2470	succinate dehydrogenase Fe-S protein subunit	108.20	3.13
CEROD2471	succinate dehydrogenase flavoprotein subunit	29.76	0.10
CEROD2472	succinate dehydrogenase hydrophobic anchor subunit	397.91	9.03
CEROD2473	Succinate dehydrogenase cytochrome b subunit	92.03	2.10
<b>cytochrome bd-type quinol oxidase</b>			
CEROD0318	Cytochrome bd-type quinol oxidase subunit 1	22.80	0.47
CEROD0319	Cytochrome bd-type quinol oxidase subunit 2	37.64	0.81
<b>NADH:ubiquinone oxidoreductase</b>			
CEROD1691	NADH:ubiquinone oxidoreductase subunit 2 (chain N)	1.45	0.00
CEROD1692	NADH:ubiquinone oxidoreductase subunit 4 (chain M)	9.10	0.37
CEROD1693	NADH:ubiquinone oxidoreductase subunit 5 (chain L)	3.65	0.64
CEROD1694	NADH:ubiquinone oxidoreductase subunit 11 (chain K)	2.81	0.00
CEROD1695	NADH:ubiquinone oxidoreductase subunit 6 (chain J)	1.99	0.00
CEROD1697	NADH:ubiquinone oxidoreductase subunit (chain I)	9.03	2.87
CEROD1698	NADH:ubiquinone oxidoreductase subunit 1 (chain H)	15.53	0.85
CEROD1699	NADH:ubiquinone oxidoreductase (chain G)	2.01	0.25
CEROD1700	NADH:ubiquinone oxidoreductase NADH-binding subunit	1.60	0.27
CEROD1701	NADH:ubiquinone oxidoreductase subunit	18.10	0.99
CEROD1702	NADH:ubiquinone oxidoreductase subunit 7	4.69	0.44
CEROD1703	NADH:ubiquinone oxidoreductase subunit	8.64	0.53
CEROD1704	NADH dehydrogenase subunit B	20.10	6.89
CEROD1705	NADH:ubiquinone oxidoreductase subunit 3 (chain A)	95.72	2.02
<b>F0F1-type ATP synthase</b>			
CEROD1756	F0F1-type ATP synthase subunit a	28.80	2.31
CEROD1757	F0F1-type ATP synthase subunit c	114.89	2.21
CEROD1758	F0F1-type ATP synthase subunit b	40.49	1.17
CEROD1759	F0F1-type ATP synthase delta subunit	8.16	1.04
CEROD1760	F0F1-type ATP synthase alpha subunit	6.97	0.24
CEROD1761	F0F1-type ATP synthase gamma subunit	0.93	0.20
CEROD1762	F0F1-type ATP synthase beta subunit	2.73	0.26
CEROD1763	F0F1-type ATP synthase epsilon subunit	1.04	0.00

**Table 10.** RPKM-values of the glycolysis, gluconeogenesis, citrate cycle and reverse citrate cycle of the epibiont in the associated state incubated in the light and in the dark with the addition of BCAAs and 2-oxoglutarate. The reference organism is *E. coli* K12 and the nearest species with suitable genome is *Chl. tepidum* TLS (Eisen *et al.* 2002).

glycolysis			
annotation	enzyme name	RPKM + light	RPKM - light
Cag_0199	glucokinase	13.67	1.74
Cag_0612	glucose-6-phosphate isomerase	14.12	1.78
Cag_1780	6-phosphofructokinase	20.81	1.16
Cag_0879	6-phosphofructokinase	1.25	0.04
Cag_0602	fructose-1,6-bisphosphatase	3.48	0.20
Cag_0728	fructose-bisphosphate aldolase	14.57	0.75
Cag_0497	triosephosphate isomerase	1.86	0.32
Cag_1420	glyceraldehyde 3-phosphate dehydrogenase	51.43	6.61
Cag_0044	phosphoglycerate kinase	8.61	0.60
Cag_0187	2,3-bisphosphoglycerate-dependent phosphoglycerate mutase	10.32	0.59
Cag_0347	enolase	16.45	0.92
Cag_2012	phosphoenolpyruvate carboxykinase	10.32	1.20
-----	pyruvate kinase		
citrate cycle and reverse citrate cycle			
annotation	enzyme name	RPKM + light	RPKM - light
Cag_2012	phosphoenolpyruvate carboxykinase	10.35	1.20
Cag_0862	malate dehydrogenase	5.11	0.08
Cag_0860	fumarate hydratase, class I	2.92	0.18
Cag_0127	succinate dehydrogenase	19.21	0.50
Cag_0128		21.10	0.67
Cag_0129		7.84	0.10
Cag_1933	succinate dehydrogenase	0.34	0.05
Cag_1934		1.53	0.25
Cag_1935		0.61	0.02
Cag_0338	succinyl-CoA synthetase	25.74	0.44
Cag_0466		3.68	0.05

Cag_0365	2-oxoglutarate ferredoxin oxidoreductase	6.01	0.11
Cag_0366		17.86	0.24
Cag_1410	isocitrate dehydrogenase	18.34	1.04
Cag_1286	aconitate hydratase 1	8.99	0.82
Cag_0317	citrate synthase	10.03	0.38
Cag_0796		48.64	1.27
Cag_0797	citrate lyase	16.85	0.41
Cag_1730	pyruvate synthase	14.14	0.37

Interestingly, after incubation in the light, the gene showing the second highest mRNA transcription of the central bacterial genome (RPKM 3439) after a hypothetical protein is the porin forming gene *ompC* (CEROD1596). A second copy of the ORF is present in the genome (CEROD1510) that is less transcribed than the first one (RPKM 128). In comparison with the incubation in the dark with BCAAs and 2-oxoglutarate, the two porins are transcribed 46.3 and 21.3 fold higher during incubation in the light.

Also highly expressed in the central bacterium incubated in the light is an  $\alpha$ -D-glucanase (CEROD1617, RPKM 562), which cleaves the  $\beta$ -glycosidic linkages of glucans releasing glucose as the sole hydrolysis product. In the dark, the transcription is lower with an RPKM value of 5.07. In addition to the  $\alpha$ -D-glucanase, the central bacterium possesses a second  $\beta$ -glucanase (CEROD1620) that is also highly transcribed with an RPKM value of 136 in the light and 4.35 in the dark. Thus, with an active epibiont,  $\beta$ -glucanases seem to be of advantage for the central bacterium.

In the dark with BCAAs and 2-oxoglutarate, only 33 genes are expressed higher as during incubation in the light. Of the 20 genes that have a predicted gene product, four are involved in chemotaxis and signal transduction. The genes CEROD1474 and 1475 encode the response regulator and the signal transduction histidine kinase of a two-component system. These systems are responsible for sensing and responding to different environmental conditions (Stock *et al.* 2000), mostly by differential expression of specific genes (Mascher *et al.* 2006). Furthermore, a PAS domain protein (CEROD0690) is differentially expressed. Those ubiquitous domains are contained in many signaling proteins where they functions as a signal sensor for various stimuli (Henry and Crosson 2011). The last of the four genes (CEROD1732) is a signal transduction protein for chemotaxis. Since the central bacterium does not seem to be metabolically active in the dark, the occurrence of four proteins involved

in sensing and responding to external stimuli hints towards an enhanced effort of the central bacterium to orientate towards more favorable conditions. However, although all the other of the 32 genes encoding a PAS domain are transcribed at least at a basic level, the bacteriophytochrome is neither transcribed in the light, nor in the dark.

In the epibiont, as expected (Tang and Blankenship 2010), the genes of both the forward and the reverse TCA are transcribed (Table 10), as well as the glycolysis. However, a pyruvate kinase has not been found in the epibiont genome. Analysis of the 12 sequenced green sulfur bacterial genomes only revealed a pyruvate kinase in the deep branching *Chloroherpeton thalassium* ATCC 35510, suggesting that in the other green sulfur bacteria, this step of the glycolysis is facilitated by a so far unknown enzyme.

In the epibiont, of the 13 genes that were expressed higher in the epibiont in the dark with organic substrates in comparison with the incubation in the light, only four had predicted proteins, a nitrogenase reductase (Cag\_1244), a redox-active disulfide protein 2 (Cag\_0399), a ferrous iron transport protein A (Cag\_1930) and a hemolysin activation/secretion protein-like (Cag\_1056). However, a symbiosis relevant characteristic of those genes during the incubation of consortia in the dark is not given. Therefore, together with the 59.1 fold increase in transcriptions for proteins assigned with translation, ribosomal structure and biogenesis for consortia incubated in the light, we conclude that the epibiont remains inactive during phases of incubation in the dark, although external carbon substrates are present.

The comparison between the epibiont in the associated and the free-living state regarding the transporters, amino acid biosynthetic pathways, glycolysis, gluconeogenesis and the forward and reverse TCA cycle revealed no significant differences (data not shown). Thus, the epibiont does not have to adapt these pathways in symbiosis, thereby confirming the postulated preadaptation of green sulfur bacteria to symbiosis (Müller and Overmann 2011; Wenter et al. 2010) and the identical results obtained from GC/MS of  $^{13}\text{C}$ -labeled amino acids in the transcriptomic comparison of free-living and associated epibionts.

## References

- Eisen, J. A., Nelson, K. E., Paulsen, I. T., Heidelberg, J. F., Wu, M., Dodson, R. J., Deboy, R., Gwinn, M. L., Nelson, W. C., Haft, D. H., Hickey, E. K., Peterson, J. D., Durkin, A. S., Kolonay, J. L., Yang, F., Holt, I., Umayam, L. A., Mason, T., Brenner, M., Shea, T. P., Parksey, D., Nierman, W. C., Feldblyum, T. V., Hansen, C. L., Craven, M. B., Radune, D., Vamathevan, J., Khouri, H., White, O., Gruber, T. M., Ketchum, K. A., Venter, J. C., Tettelin, H., Bryant, D. A., and Fraser, C. M. (2002). The complete genome sequence of *Chlorobium tepidum* TLS, a photosynthetic, anaerobic, green-sulfur bacterium. *Proceedings of the National Academy of Sciences of the United States of America* 99(14), 9509-9514.
- Henry, J. T., and Crosson, S. (2011). Ligand-binding PAS domains in a genomic, cellular, and structural context. *Annu Rev Microbiol* 65, 261-86.
- Imhoff, J. F. (2003). Phylogenetic taxonomy of the family Chlorobiaceae on the basis of 16S rRNA and fmo (Fenna-Matthews-Olson protein) gene sequences. *International Journal of Systematic and Evolutionary Microbiology* 53(4), 941-951.
- Mascher, T., Helmann, J. D., and Uuden, G. (2006). Stimulus perception in bacterial signal-transducing histidine kinases. *Microbiol Mol Biol Rev* 70(4), 910-38.
- Möglich, A., Ayers, R. A., and Moffat, K. (2009). Structure and signaling mechanism of Per-ARNT-Sim domains. *Structure* 17(10), 1282-94.
- Müller, J., and Overmann, J. (2011). Close interspecies interactions between prokaryotes from sulfurous environments. *Frontiers in Microbiology* 2.
- Overmann, J., and Schubert, K. (2002). Phototrophic consortia: Model systems for symbiotic interrelations between prokaryotes. *Archives of microbiology* 177(3), 201-208.
- Stock, A. M., Robinson, V. L., and Goudreau, P. N. (2000). Two-component signal transduction. *Annu Rev Biochem* 69, 183-215.
- Tang, K. H., and Blankenship, R. E. (2010). Both forward and reverse TCA cycles operate in green sulfur bacteria. *Journal of Biological Chemistry* 285(46), 35848-35854.
- Tatusov, R. L., Galperin, M. Y., Natale, D. A., and Koonin, E. V. (2000). The COG database: a tool for genome-scale analysis of protein functions and evolution. *Nucleic Acids Res* 28(1), 33-6.
- Wenter, R., H • tz, K., Dibbern, D., Li, T., Reisinger, V., Ploscher, M., Eichacker, L., Eddie, B., Hanson, T., Bryant, D. A., and Overmann, J. (2010). Expression-based



identification of genetic determinants of the bacterial symbiosis ' *Chlorochromatium aggregatum* '. *Environmental microbiology*.

Yadav, V. K., and Archer, D. B. (1989). Sodium molybdate inhibits sulphate reduction in the anaerobic treatment of high-sulphate molasses wastewater. *Applied Microbiology and Biotechnology* 31(1), 103-106.



## Chapter 5

### Discussion

#### 5.1 Phylogeny of the green sulfur bacteria

The presented work represents the most comprehensive one on the family of the *Chlorobiaceae* conducted so far. As shown by rarefaction analysis, it comprises the majority of the existing green sulfur bacterial diversity. But the question arises, if the primers used to gather the sequences utilized in this study detected the whole diversity of the GSB or if a part of the diversity was missed because the primers were too specific. The currently used primers for GSB have originally been based on only 13 sequences available (Overmann and Tuschak 1997; Tuschak et al. 1999) and might therefore be too specific to capture all green sulfur bacterial sequences. This concern, however, can be dismissed when looking at the primers used to amplify the 16S rDNA that were used to calculate the respective tree. Of the 107 studies that contributed sequences for the calculation of the tree, 64 studies have been using universal primers for bacterial 16S rDNA, while only 12 studies used primers specific for the amplification of green sulfur bacterial 16S rDNA. In the remaining studies, the 16S rDNA of cultivated GSB was sequenced. Thus, the majority of the sequences was collected using universal 16S rDNA primers. We therefore assume that no important groups of GSB have been overlooked due to the specificity of the green sulfur bacterial primers.

Besides the specificity of the primers, the global distribution of sampling sites might play a role in covering the green sulfur bacterial diversity. So far, in large areas of the earth, no GSB have been detected (Fig. 6). From the inland of Africa, South America and Australia, no GSB sequences are available, which is most probably due to a lack of investigations of meromictic lakes in those areas, since no literature on this topic can be found. Only a few marine studies have provided green sulfur bacterial sequences in those regions. Thus, with the majority of the sampling sites concentrated in Europe, North America and Asia, nearly all of the sequences obtained so far descend from the northern hemisphere of the earth. The question therefore arises, if part of the green sulfur bacterial diversity is missed due to a lack of sampling in the southern hemisphere.

A first answer comes from the few green sulfur bacterial sequences obtained from the southern hemisphere. They are similar or even identical to already available sequences. The three sequences obtained from Ace Lake, Antarctica cluster in group 2a, with one of the sequences being identical to *Chl. luteolum* DSM 270<sup>T</sup> found in Lake Polden, Norway. And the sequence of *Chl. tepidum* ATCC 49652 obtained from New Zealand, clusters with environmental sequences from Denmark and Spain. Thus, the geographical distance does not seem to limit the distribution of these strains. Biogeographical studies have revealed, that many prokaryotes have a cosmopolitan distribution in their respective habitats up to the genus level (Hedlund and Staley 2003). Furthermore, at the level of 16S rRNA gene resolution, identical phylotypes can be identified in similar habitats in different geographic areas. This holds true for marine (Massana et al. 2000) as well as for fresh-water (Glöckner et al. 2000) prokaryotes. In the case of GSB, all four phylogenetic groups are found on the North American, the European and the Asian continent respectively. A restriction of green sulfur bacterial groups to certain continents is unapparent. It is however noticeable, that in the well sampled countries India and Germany, certain groups of GSB are predominantly found. In India, in only two of the 20 sampling sites investigated group 3 GSB were found, whereas in the rest of the sampling sites sequences from group 1 and 2 were found exclusively. In Germany, sequences from group 2 and 3 were dominant. Those sequences suggest, that the distribution of the phylotypes they derived from is restricted by geographical boundaries. Such a non-cosmopolitan distribution has also been found for the cyanobacterial genus *Synechococcus*, which was identified in hot springs in North America but not in other hot springs around the globe. It is suggested that physical isolation and a lack of viability during transport are restricting their dispersal (Castenholz 1978; Papke et al. 2003). For the obligate anaerobic GSB, especially for fresh-water strains that are colonializing non-connected water-bodies, the transport over longer distances is supposedly a problem as well. However, identical 16S rDNA sequences from fresh-water strains are often found on two or even three continents. Thus, also strictly anaerobic, certain types of GSB seem to be able to cross longer distances without losing their viability.

To assess whether the majority of the green sulfur bacterial diversity is comprised by this study is of interest when discussing the results of the rarefaction and species richness analysis. If a group of GSB is overlooked because of primers that are too specific to detect the existing GSB or insufficient sampling, the analysis would be biased. The effects of such a bias can be observed when looking at the rarefaction analysis of the cultivated GSB on the

97% sequence similarity level. The ascend of the slope is less steep than that of the full set of the sequences. This means, that from certain clusters of the phylogenetic tree, only few or no isolates have been obtained. The reason for this is very likely the utilization of culture medium that is selective for the majority of strains, but is insufficient to grow a few so far uncultured ones. Of the 51.7 green sulfur bacterial strains predicted to exist by estimateS, presumably only 39.8 are able to be cultivated with the culture media used so far. This means that approximately 23% of the green sulfur bacterial diversity is not going to be cultivated using to days cultivation methods. Additionally, the species missed might be of special interest since they require different cultivation methods and therefore differ from the species cultivated to date.

Especially in subgroup 2a and 3a, only few strains have been isolated so far. In group 2a, no isolates have been obtained from a cluster of 5 sequences and in a cluster of 11 sequences, only *Chlorobium luteolum* E1P1 is cultured although 4 different species are predicted from distance matrix comparison. In subgroup 3a, a representative in a cluster of 6 sequences consisting only of environmental sequences is missing. Also, from a cluster in group 4a consisting of 12 environmental sequences found in four different sampling sites in India, a cultured species has not been obtained so far. These examples confirm the prediction by estimateS that cultivated species are underrepresented. Thus, some green sulfur bacterial strains seem to be more adapted to the conditions in their respective habitat and are therefore not culturable under the standard conditions used for GSB. Experimenting with the culture medium and incubation conditions might be necessary to isolate these so far uncultured GSB.

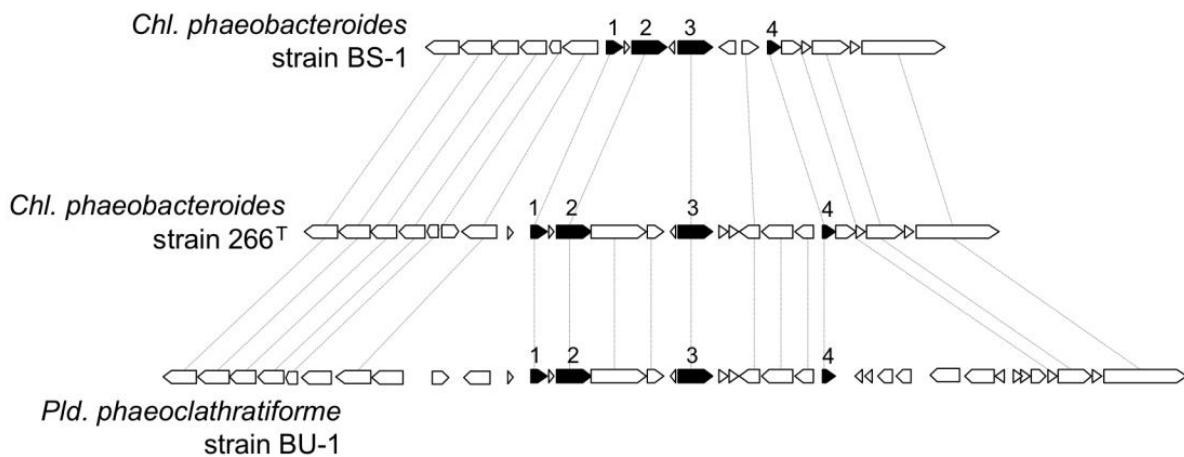
### 5.1.1 Diversification of green sulfur bacteria

The family of the *Chlorobiaceae* is a phylogenetically shallow group with high similarity values among 16S rDNA gene sequences ( $>90.1\%/K_{\text{nuc}} < 0.11$  (Overmann 2000)), with the sole exception being *Chloroherpeton thalassium* (85.5–87%) (Overmann 2001). However, within this family, a large diversity is found. But the question arises, which selection factors drive the diversification of the family *Chlorobiaceae*. Since their habitat is limited to the chemocline of meromictic lakes and water bodies, variables that lead to niche formation are equally limited as well.

One factor that contributes to the diversification of GSB is light. Depending on the depth of the chemocline, the turbidity or the pigmentation of organisms in the water layers

above the chemocline, light of different wavelength reaches the habitat of the GSB. Depending on the wavelength of the light reaching the chemocline, either green- or brown-colored types of GSB are found predominantly. The green colored types of GSB thrive in dystrophic lakes (Parkin and Brock 1980) or underneath mats of purple sulfur bacteria, where light of the blue or red wavelength range prevails (Caldwell and Tiedje 1975; Gorlenko and Kuznetsov 1971; Overmann et al. 1998a). Further distinguishing the green types of GSB is their ratio between bacteriochlorophyll (bch) *c* and *d*. Strains with high amounts of bch *c* grow faster under low-light conditions than mutants of the same strain only being able to produce bch *d* (Maresca et al. 2004).

The brown types, on the other hand, are primarily found in eutrophic lakes or at greater depth, where only light in the blue-green to green wavelength range is available. Under these conditions, brown-colored types of GSB containing high concentrations of the light-harvesting carotenoids isorenieratene and  $\beta$ -isorenieratene instead of chlorobactene, have a selective advantage over their green-colored counterparts (Montesinos et al. 1983). When looking at the distribution of green- and brown-colored strains among the 16S rDNA tree of GSB, clusters of green or brown types cannot be identified. The only exception is the absence of brown strains in subgroup 4a. Albeit, only of 5 of the 14 cultured strains in this subgroup, information about their pigmentation is available. The relatively even distribution of green and brown strains in the phylogenetic tree suggests that differing wavelength ranges are not a major driving force in the diversification of GSB. This is supported by the occurrence of strains with identical 16S rDNA sequences that differ in their particular color. For a green-colored GSB to become a brown-colored one, the monocyclic chlorobactene needs to be transformed to the dicyclic  $\beta$ -isorenieratene and bch *c* has to be formylated at the C7 position to be converted to bch *e* (Brockmann 1976). Initiating the transformation of chlorobactene to  $\beta$ -isorenieratene is the gene *cruB*, which is exclusively found in brown-colored GSB. In the genomes of the three sequenced brown-colored GSB species, it lies in a small gene cluster that includes three other genes, one of which is a candidate for the conversion of bch *c* to *e* (Figure 20.) (Maresca 2007).



**Figure 20.** Gene neighborhoods around *cruB* (#3) in *Chl. phaeobacteroides* BS-1, *Chl. phaeobacteroides* DSM 266<sup>T</sup>, and *Pld. phaeoclathratiforme* BU-1. #1. 2-vinyl bacteriochlorophyllide hydratase (*bchF*), #2. Hypothetical protein (radical SAM motif), #4. Isoprenylcysteine carboxyl methyltransferase (Maresca 2007).

It is therefore speculated that this cluster is easily transferred between green sulfur bacterial species and responsible for the conversion from a green- to a brown-colored strain. The even distribution of the two types of GSB we found among the 16S rDNA tree supports the hypothesis of lateral gene-transfer as reason for the switch from green- to brown-colored strains.

Interestingly, all but one green sulfur bacterial sequence found on corals form a distinct cluster in the marine group, with *Prosthecochloris vibrioformis* 3M being the only cultured representative. Very likely, this type of green sulfur bacterium is part of a bacterial community responsible for coral diseases. Of the five studies that contributed green sulfur bacterial sequences associated with corals, four were investigating diseased tissue (de Castro et al. 2010; Frias-Lopez et al. 2002; Myers and Richardson 2009; Pantos et al. 2003). The black band disease dissolves corals by creating a sulfide-rich microenvironment (Richardson et al. 2009). This explains why GSB are frequently found associated with corals, which are usually not located in the chemocline but in the oxic zone of the sea. The fact that all sequences were clustered together suggests that a genetic adaptation was necessary to colonize this specific niche.

From the temperature of the habitat, a connection can be drawn from the physiology to the phylogeny. Most sequences derived from warm or hot habitats are found in group 4. But thermophilic GSB are not found in this cluster exclusively but also in group 4b and 3b.

Therefore, this trait is either polyphyletic or was acquired by other strains through horizontal gene transfer. This is similar to the trait of salt tolerance in GSB. Until recently, the salt requirements of GSB were believed to be a reliable criterion to differentiate between green sulfur bacterial species (Alexander et al. 2002; Imhoff 2001; Overmann 2000) and therefore are a result of niche formation within the GSB. However, as it has been shown lately (Triado-Margarit et al. 2010), salt tolerance does not seem to be restricted to species of any specific group, but is widespread among the phylogeny of GSB, and closely-related phylotypes can have dissimilar salt tolerance capacities. Thus, also salt tolerance most likely lead to the formation of the marine group, it is not a criterion that is strictly limited to one group of green sulfur bacteria. Since lateral gene transfer is responsible for the color-switch from green to brown pigmentation, one could speculate that GSB frequently exchange genetic material that leads to an extensive distribution of physiological properties. This frequent exchange would also explain the futility of physiological traits that have been used to characterise the family of the *Chlorobiaceae* like pigmentation, motility, cell morphology and the ability to form gas vesicles (Pfennig and Trüper 1989).

### 5.1.2 Analysis of the marker genes and the ITS region

The *fmoA*, *bchg* and *sigA* gene sequences as well as the ITS region have been used in phylogenetic studies on eubacteria and GSB. By distance matrix comparison, the highest correlation to the 16S rDNA was found for the FMO protein, which is congruent with the finding of Imhoff (Imhoff 2003) and our own finding, that the *fmoA* gene sequence tree shows a high correlation to the 16S rDNA tree (Suppl. Fig. S1). A lesser degree of similarity was found for the *bchg* and *sigA* gene sequences. Thus, from the results of the distance matrix comparison, the *fmoA* gene sequence is best suited to support the phylogenetic classification of the 16S rDNA in green sulfur bacteria.

As it has been expected, the distance matrix comparisons for the coding bases of the gene sequences show a higher correlation to the 16S rDNA than the complete sequence. Therefore, the concatenated tree was calculated using the first two bases of the three marker genes. The resulting tree nearly matches the 16S rDNA tree, proposing that the functional genes investigated have derived from a common ancestor and have not been exchanged by lateral gene transfer.



The high similarity of the concatenated tree to the 16S rDNA tree was expected from the distance matrix comparison values for the first two bases. With an average of 0.8, the correlation coefficient of the first two bases of the concatenated sequences is higher than the correlation coefficient of the complete *fmoA* gene sequence that already depicts the phylogeny of the GSB similarly to the 16S rDNA sequence.

One difference between the concatenated and the 16S rDNA tree is the classification of *Chlorobium chlorochromatii* CaD3<sup>T</sup>, the epibiont of the phototrophic consortium “*Chlorochromatium aggregatum*”. It cannot be assigned to any of the four groups of GSB by 16S rDNA phylogeny. But, as already proposed by Imhoff based on the *fmoA* gene sequence (Imhoff and Thiel 2010), the concatenated sequence phylogeny confirms the classification of strain CaD3<sup>T</sup> as a member of the genus *chlorobium*.

Another inconsistency between the 16S rDNA and the concatenated tree is that *Cba. limnaeum* UdG 6045, 6040 and 6041 have identical concatenated but not identical 16S rDNA sequences. This however is easily explained by looking at the 16S rDNA sequences. These contain several ‘N’s where the correct nucleotides could not be assigned, which speaks for an overall flawed sequencing of these strains. Therefore, the sequences are not identical even when leaving out the ‘N’s. Thus, in this case, where the marker genes are identical and the sequencing of the 16S rDNA is obviously flawed, it is very likely that the 16S rDNA sequences of these strains are identical or nearly identical to each other as well.

With the exception of *Chl. chlorochromatii*, the sequences in the concatenated tree cluster in the same groups than the 16S rDNA sequences. With three conserved marker genes used, the concatenated tree is an accurate analysis confirming the current phylogeny of the GSB. In addition, the distances in the concatenated tree are longer than in the 16S rDNA tree, so that the concatenated tree provides a higher phylogenetic resolution. However, conducting this analysis is complex and using only one marker gene might be sufficient to get similar results. We therefore investigated which of the marker genes used has the most potential to be utilized as an additional marker gene besides the 16S rDNA. As already described, the *fmoA* gene sequence has the highest correlation to the 16S rDNA in the distance matrix comparison. However, if only the first two bases would be used, *bchG* would yield nearly similar results. But the *fmoA* gene sequence not only shows the highest similarity, but also the highest phylogenetic resolution. In comparison with the *bchG* and *sigA* gene sequence, *fmoA* has the highest pair-wise sequence dissimilarity when looking at the complete sequence and the coding bases. Thus, although correlating more closely to the 16S rDNA, the *fmoA* sequences

differ more from each other than the *sigA* and *bchG* gene sequences and therefore allow for a deeper resolution of the green sulfur bacterial phylogeny. The higher pair-wise sequence dissimilarity of *fmoA* is rooted in more frequent mutations at non-synonymous sites of the gene, which means at sites where a point-mutation leads to a change of the coded amino acid. All three genes investigated are exposed to a strong purifying selection pressure, which means that most mutations in non-synonymous positions are negative and lead to extinction by selection. This is very likely due to all three genes being essential to the viability of the organism. The *fmoA* gene sequence however is less exposed to selection pressure as the other two genes investigated and seems to underlie a relaxed purifying pressure. As has been shown by the alignment of 17 *fmoA* gene sequences (Larson et al. 2011), the FMO protein has highly conserved regions at the binding sites of the chlorophyll *a* molecules and at the sites of interaction with other proteins. Thus, the non-synonymous mutations have to be located at the less conserved non-binding and non-interacting sites that are less crucial for the correct functioning of the protein. These mutations are responsible for the higher sequence dissimilarity of the *fmoA* gene sequence and enhance its resolution of the green sulfur bacterial diversity.

The ITS region did not show any significant correlation to the 16S rDNA. It might therefore be used to further distinguish strains that are identical in their 16S rDNA and in the marker genes used in this study. However, green sulfur bacterial strains that are identical in their 16S rDNA and the three marker genes utilized in this study, for example the seven identical strains of group 3b depicted in the concatenated tree, cannot be further distinguished using the ITS region (Suppl. Fig. S4). Another example are the strains *Cba. limnaeum* UdG 6045, 6040 and 6041. The three marker genes and the ITS region of those strains are identical. Their 16S rDNA differs slightly, this however is very likely due to a flawed sequencing (s.discussion above). From the sequence information available, those three strains are identical although they have been isolated from two different lakes in Spain and one lake in the USA. The large-scale geographical distribution and the identical sequences of these strains hint towards a recent dispersion of these strains.

### 5.1.3 Future sequencing of green sulfur bacterial strains

From the family Chlorobiaceae, 12 strains have been sequenced so far. With the phylogenetic tree calculated in this study, interesting targets for further genome sequencing can be sought

out. A possible candidate is *Prosthecochloris vibrioformis* 3M, the only cultured representative in the cluster of coral-associated GSB. It is not only distantly related to the other two sequenced strains of the marine group, *Prosthecochloris vibrioformis* DSM 271<sup>T</sup> and *Chlorobium phaeobacteroides* BS-1, but might also provide insight in why the sequences clustered around this strain are found exclusively on saltwater corals.

Another interesting target for genome sequencing might be *Prosthecochloris vibrioformis* DSM 260<sup>T</sup>. It is the type strain of a distinct phylogenetic cluster encompassing eleven cultivated strains and could therefore provide information on a number of already cultured GSB. Quite the opposite, *Chlorobium luteolum* E1P1 is the only isolate in a distinct cluster consisting of eleven sequences and could therefore provide information on the nature of this subgroup of GSB.

Although the mentioned strains might provide additional information about the physiology and phylogeny of GSB, the so far sequenced genomes cover the rough diversity of GSB. From each group or subgroup of GSB, at least one complete genome is available. In addition, the only cultured epibiont of a phototrophic consortium *Chlorobium chlorochromatii* CaD3 and the more distantly related *Chloroherpeton thalassium* ATCC 35110<sup>T</sup> has been sequenced. Therefore, genome comparison using the available green sulfur bacterial genomes (Davenport et al. 2010) already encompass the borders of the green sulfur bacterial phylogeny, thus providing information from which an understanding on other green sulfur bacterial strains can be derived.

## 5.2 Molecular basis of the symbiosis in the phototrophic consortium “*C. aggregatum*”

### 5.2.1 Transfer of metabolites from the epibiont to the central bacterium

Although phototrophic consortia have been known since 1906 (Lauterborn 1906), the reasons why green sulfur bacteria and heterotrophic *Betaproteobacteria* form these intimate associations are only partially known to date. The chemotaxis of consortia towards light and sulfide (Fröstl and Overmann 1998) explains how the epibiont profits from this partnership, but leaves the question of advantages for the central bacterium unacknowledged. One characteristic feature of green sulfur bacteria that provides a basis for interaction with other bacteria is the extracellular deposition of sulfur globules (zero valence sulfur), the initial

product of sulfide oxidation during anoxygenic photosynthesis. This sulfur is further oxidized to sulfate only after depletion of sulfide. The extracellular deposition renders the sulfur available to other bacteria such as, e.g., sulfur reducers. Therefore, it had initially been proposed that the central bacterium of phototrophic consortia is a sulfate- or sulfur-reducing bacterium. In that case, extracellular sulfur produced by the green sulfur bacteria could be utilized by the central bacterium to establish a close sulfur cycle within the phototrophic consortium (Pfennig 1980). Such a sulfur cycle has been established in defined syntrophic cocultures of *Chlorobium phaeovibrioides* and *Desulfuromonas acetoxidans*. In these co-cultures, acetate is oxidized by *Desulfuromonas acetoxidans* with sulfur as electron acceptor, which leads to a recycling of the sulfide that can then be used again for anoxygenic photosynthesis by *Chl. phaeovibrioides*. Only minute amounts of sulfide (10  $\mu\text{M}$ ) are required to keep this sulfur cycle running (Warthmann et al. 1992). Similarly, sulfate reducers are able to grow syntrophically with green sulfur bacteria with only low equilibrium concentrations of sulfide (Biebl and Pfennig 1978). In addition, such interactions may also encompass transfer of organic carbon compounds between the partners. In mixed cultures of *Desulfovibrio desulfuricans* or *D. gigas* with *Chlorobium limicola* strain 9330, ethanol is oxidized to acetate with sulfate as electron acceptor and the acetate formed is incorporated by *Chl. limicola* such that ethanol is completely converted to cell material.

However, the hypothesis of a sulfur cycling within phototrophic consortia became less likely by the discovery that the central bacterium belongs to the *Betaproteobacteria*, whereas only the *Deltaproteobacteria* or *Firmicutes* encompass typical sulfur- or sulfate-reducers (Fröstl and Overmann 2000). The genome sequencing of the central bacterium showed the central bacterium to be incapable of sulfur respiration. Therefore, the possibility of a sulfur cycle within “*C. aggregatum*” recycling sulfate to sulfide can now be ruled out.

Instead of a sulfur cycle, the physiological experiments described in this thesis demonstrate that carbon is transferred from the epibiont to the central bacterium. Autotrophy and thus fixation of  $\text{HCO}_3^-$ , which was used in NanoSIMS,  $^{14}\text{C}$ - and  $^{13}\text{C}$ -labeling experiments, can only be performed by the epibiont but not by the central bacterium according to genome analysis and negative controls in the dark. Furthermore, no other photosynthetic organisms are present in the consortia enrichment culture (data not shown) and the incubation in the dark revealed no incorporation of labeled carbon into the cells. Therefore,  $\text{HCO}_3^-$  has to be fixed by the epibiont and afterwards transferred to its symbiotic partner. This transfer can already be detected after 15 min of incubation with  $\text{H}^{13}\text{CO}_3^-$ , at which point the label in the central

bacterium is already similar to the one in the epibiont. Since no time delay between the incorporation of labeled carbon into the epibiont and the central bacterium occurred, we conclude that newly synthesized small molecular weight organic matter is transferred rather than macromolecules.

The nearly identical  $^{13}\text{C}$ -labeling patterns of amino acids in the epibiont and the central bacterium incubated with  $\text{H}^{13}\text{CO}_3^-$  suggest, that one group of transferred organic matter is amino acids. As the analysis of the epibiont pure culture supernatant revealed, epibionts indeed excrete substantial amounts of amino acids. Particularly high is the excretion of dissolved combined amino acids and dissolved free BCAAs. Both forms of amino acids could be excreted by an ABC-type oligopeptide transporter and an ABC-type branched chain amino acid transporter. Both transporters are highly transcribed in the associated epibiont incubated in the light. In natural aquatic habitats, concentrations of DCAA are generally higher than concentrations of DFAA (Coffin 1989; Kroer *et al.* 1994) and can account for up to 45 % of the carbon and 112 % of the nitrogen requirements in the heterotrophic bacterial ecosystem (Kroer *et al.* 1994). It is therefore an important substrate in natural habitats. The central bacterium is capable of taking up DCAA as well as BCAAs via the two ABC-type oligopeptide transporter systems that are encoded in the genome and the two ABC-type branched chain amino acid transporter systems. Of each transporter type, one is transcribed under standard cultivation conditions. With the addition of externally added  $[\text{U}-^{13}\text{C}]$  leucine, we were able to show that the central bacterium is indeed taking up and incorporating external BCAAs.

In summary, our investigations showed that (i) the central bacterium is able to take up and incorporate externally added  $[\text{U}-^{13}\text{C}]$  leucine (ii) the epibiont is excreting high amounts of DCAA and BCAAs, (iii) the amino acid labeling patterns in the two partner organisms is nearly identical and (iiii) transcribed oligopeptide- and BCAA transporters encoded in the genome are transcribed in both symbiotic partners. Based on this combined evidence, we conclude that a transfer of amino acids mainly in the form of DCAA and BCAAs conducted via transporters occurs from the epibiont to the central bacterium. The transfer of amino acids has been shown to be of importance in other close prokaryotic interactions as well. In the archaea-archaea association of *Ignicoccus hospitalis* and *Nanoarchaeum equitans*, identical amino acid labeling patterns and the lack of amino acid biosynthesis pathways in *N. equitans* demonstrated an exchange of amino acids in this tight association (Jahn *et al.* 2008). In an analogous example, amino acids are used as both carbon and ammonium shuttles in legume-

root nodules (Lodwig *et al.* 2003). Especially the property of amino acids of providing a carbon and nitrogen source is interesting in respect to the symbiosis of phototrophic consortia. Although the K3 cultivation medium provides  $\text{NH}_4^+$  as nitrogen source that can be used by both the epibiont and the central bacterium, in the natural aquatic habitat, nitrogen limitation is frequently observed (Kosten *et al.* 2009; May *et al.* 2010; Migal 2011; Wu *et al.* 2006). Being associated with a green sulfur bacterium able to perform nitrogen fixation and transferring the fixed nitrogen via amino acids is therefore providing the central bacteria with a competitive advantage over their heterotrophic competitors.

In addition to nearly identical labeling patterns, the quantification of the labeled amino acids in the central bacterium and the epibiont after one hour of incubation revealed a similar distribution of  $^{13}\text{C}$ -label among the amino acids. However, serine, aspartate and especially glutamate differed. Glutamate is the most highly labeled amino acid in the epibiont. Usually, high concentrations of glutamate are common in growing bacteria as it serves as an amino-shuttle during the synthesis of other amino-acids (Tyler 1978). Interestingly, this amino acid that is highest labeled and needed to trans-aminated amino acids in the epibiont is not transferred to the central bacterium after 1 h of incubation and low concentrations of labeled glutamate can only be found after 3 h of incubation with  $\text{NaH}^{13}\text{CO}_3$ . This suggests a non-random, selective excretion of amino acids attuned to the needs of the epibiont cells.

As opposed to the lack of amino acid synthesis genes in *N. equitans* (Jahn *et al.* 2008), all genes necessary for the synthesis of the 20 proteogenic amino acids except *serA* are present in the genome and mostly transcribed in the central bacterium. Thus, the heterotrophic partner organism does not solely rely on the transfer of amino acids by the epibiont to synthesize proteins as it has been observed in rhizobia that become symbiotic auxotrophs for branched chain amino acids after infecting the host plant due to a downregulation of the respective biosynthetic pathways (Prell *et al.* 2009). The hypothesis that the central bacterium is additionally synthesizing its own amino acids is also supported by the increasing dissimilarity between the labeling patterns of the two partner organisms observed between 1 h of incubation and 3 h of incubation with  $\text{NaH}^{13}\text{CO}_3$  (Fig. 13). The transcribed genes for amino acid synthesis and the growing dissimilarity of the amino acid labeling patterns suggest that organic substrates are available to the central bacterium which can be used for *de-novo* synthesis of amino acids and thereby account for the differences in the labeling patterns. Several candidates can be identified as substrates. Firstly, the transferred amino acids themselves could be metabolized to synthesize other amino acids. Secondly, keto acids like 2-

oxoglutarate which are also excreted by green sulfur bacteria (Sirevag and Ormerod 1970) could be taken up by the central bacterium. And thirdly, green sulfur bacteria are able to synthesize glycogen as non-membrane-bound granules, usually in the presence of excess carbon substrates and light energy but limiting amounts of inorganic nutrients like ammonium and phosphate (Sirevag and Ormerod 1977). Of the three enzymes necessary for glycogen synthesis (Preiss 1984), a glycogen synthase is present in the epibiont genome and transcribed during cultivation in the light (RPKM 7.38). For the other two enzymes of glycogen synthesis, the branching enzyme and the ADP-glucose pyrophosphorylase, no direct hit was found in the genome annotation. But these two enzymes are also missing in the completely sequenced green sulfur bacterial strain *Chlorobaculum parvum* NCIB 8327, in which glycogen storage was experimentally verified (Sirevag and Ormerod 1977). Thus, an alternative pathway is used to synthesize glycogen in *Chl. parvum* which might also be used by the epibiont. This is further endorsed by the observation that the stained glycogen globules in *Cba. parvum* highly resemble the structures observed in the epibiont in trans-electron-microscope pictures of consortia after cryofixation by high-pressure freezing (Wanner et al. 2008). In addition to this finding, a highly transcribed permease (Cag\_1339) involved in carbohydrate transport is found in the epibiont and four proteins responsible for sugar transport in the central bacterium (CEROD0679, CEROD0555, CEROD1125, CEROD0642) were identified. Furthermore, an  $\alpha$ -D-glucanase (CEROD1617) and a  $\beta$ -glucanase (CEROD1620) that cleave the  $\beta$ -glycosidic linkages of glucans releasing glucose as the sole hydrolysis product are highly transcribed in the central bacterium in the light. In the epibiont, a cellulose synthase that, similar to the glycogen synthase, is using UDP-glucose as substrate for chain-elongation, is encoded in the genome and transcribed at a basic level in the light. Based on this data, we conclude that besides amino acids, sugar is another group of substrates that is transferred within the phototrophic consortium. The exchange of carbon is well investigated in arbuscular mycorrhiza, the mutualistic symbiosis between a fungus and the roots of a plant (Kirk et al. 2001), in which phosphate provided by the fungus is exchanged against carbon supplied by the plant (Harrison 2005). The interaction is highly developed, with partners offering the best rate of exchange being rewarded, thereby creating an evolutionary stable mutualistic interaction (Kiers et al. 2011). In the same manner, central bacteria moving the epibiont into the light and sulfide ensure their own supply with substrates provided by the photosynthetic partner.

In the presence of external substrates however, the transfer between the epibiont and the central bacterium is reduced. The addition of BCAAs and 2-oxoglutarate both resulted in a reduction of radiolabel transfer from the epibiont to the central bacterium. Thus, external carbon sources are incorporated into the central bacterium in addition to the metabolites provided by the epibiont. Besides 2-oxoglutarate and the combination the three BCAAs together, arginine, glutamine and leucine have shown to reduce the carbon transfer from the epibiont to the central bacterium. Therefore, the central bacterium is not solely relying on the epibiont to provide organic substrates, but is taking up substrates for its metabolism from the surrounding medium. This is in line with the observation that the inhibiting effect on the transfer between the two partners by external substrates is diminished when a higher  $\text{NaHCO}_3$  concentration is used during incubation. With 1 mM of either  $[\text{U-}^{13}\text{C}]$  leucine or  $[\text{U-}^{13}\text{C}]$  2-oxoglutarate but 10 mM of  $\text{NaHCO}_3$  in the culture medium instead of 30  $\mu\text{M}$ , only a fraction of the amino acids were labeled.  $[\text{U-}^{13}\text{C}]$  leucine was taken up by both the central bacterium and the epibiont. In the control essay with 10 mM  $\text{NaH}^{13}\text{CO}_3$  as only carbon source,  $9.97 \pm 3.91$  % of the carbon in the central bacterium amino acids were labeled after 3 h. Thus, in the excess of  $\text{NaHCO}_3$ , the majority of carbon is provided by the epibiont as opposed to the incubation of consortia with a lower  $\text{NaHCO}_3$  concentration, whereby external substrates are predominantly used by the central bacterium.

### 5.2.2 Mechanisms of metabolite exchange

This scheme of a competitive uptake of substrates according to their availability supports the hypothesis that the central bacterium is taking up its substrates, whether derived from the epibiont or externally, via a variety of transporters. Several observations confirm this assumption.

At first, the very structure of the phototrophic consortium in itself facilitates the putative transfer of compounds from one partner to the other since the direct cell-cell-contact prevents a diffusion of compounds over larger distances and hence minimizes transfer time. In that regard, the ultrastructure of the contact sites in "*C. aggregatum*" has been studied in detail using different electron microscopy approaches (Wanner et al. 2008). In free-living epibionts, as well as in all other known green sulfur bacteria, chlorosomes are distributed evenly among the inner face of the cytoplasmic membrane. However, in the associated state, chlorosomes are absent in the green sulfur bacterial epibionts at the site of attachment to the central



bacterium. Replacing the antenna structures, a 17 nm-thick layered structure of yet unknown function has been discovered (Vogl et al. 2006; Wanner et al. 2008). Interestingly, treatment of the epibionts with the extracellular cross-linkers DTSSP and BS<sub>3</sub> revealed the branched chain amino acid ABC-transporter binding protein (compare section 5.2) to be cross-linking with other proteins, indicating that it is localized at the cell surface or in the periplasm (Wenter et al. 2010).

In "*C. aggregatum*", prominent connections between the two partner bacteria are stretching out from the central bacterium to the epibiont. Those periplasmic tubules (PT) are formed by the outer membrane and are in linear contact with the epibionts. The PT are distributed over the entire cell surface and reach 200 nm in length at the poles of the central bacterium. It had been speculated that the periplasmic tubules represent connections of a shared periplasmic space (Wanner et al. 2008). However, this could not be confirmed by Fluorescence-Recovery-After-Photobleaching (FRAP)-analysis. After staining of the consortia with *calcein* acetoxymethylester (*calcein AM*), only the epibionts but not the central bacterium could be detected by fluorescence microscopy. This result in itself already contradicts the hypothesis of a combined periplasm because the highly fluorescent dye *calcein* should have diffused after its formation from the epibiont cells into the central bacterium. Furthermore, after subsequent bleaching of one of the epibiont cells using confocal microscopy, a recovery of fluorescence could not be detected in the bleached cell, which excludes the possibility of free diffusion between the epibiont cells through the interconnecting pili. A similar experiment has been conducted with the filamentous cyanobacterium *Anabaena cylindrica* which is considered to be a truly multicellular prokaryote. Single *calcein* stained cells within an *Anabaena* filament fully recovered *calcein* fluorescence twelve seconds after bleaching. This effect was ascribed to intercellular channels allowing free diffusion of molecules from cytoplasm to cytoplasm (Mullineaux et al. 2008). Such a rather unspecific transfer is unlikely to occur across the contact site of the phototrophic consortium "*C. aggregatum*" where it must be much more substrate-specific.

After the possibility of free diffusion within the cells of the consortium has been ruled out, the question remains why the central bacterium is not stained after treatment with *calcein-AM*. Either, the central bacterium is lacking an esterase specific for cleaving *calcein-AM*, or the tight packing of epibiont cells represent a diffusion barrier around the central bacterium preventing dissolved compounds to diffuse into the consortium and reach the central bacterium. This question was addressed by adding carboxyfluorescein diacetate

succinimidyl ester (CFDA-SE) to intact consortia and following fluorescence in epibionts and central bacteria over time (Bayer 2007). CFDA-SE enters cells by diffusion and, after cleavage by intracellular esterase enzymes, confers fluorescence to the bacterial cell. Central bacteria were already detectable after two minutes of exposure, whereas epibionts in the same sample could only be detected after 12 minutes of incubation with CFDA-SE and developed only weak fluorescence. The fluorescence activity of the central rod remained strong throughout the experiment which is indicative of a higher esterase activity than in the epibiont. Since the central bacterium is presumably heterotrophic, it is likely to express esterases such as lipases at a higher number or intracellular concentration. These results indicate that, even in intact phototrophic consortia, diffusion of small water-soluble molecules towards the central bacterium is not significantly impeded by the surrounding layer of epibionts. This is in line with the hypothesis that the uptake of molecules derived from the epibiont is mediated by transporters of the central bacterium.

In addition to transporters, the OmpC porin might play an important role in substrate uptake. The *ompC* gene is present in two copies in the central bacterium, which are both highly transcribed, with CEROD1596 even being the second highest transcribed gene in the central bacterium. Trimers of OmpC form a 1.3 nm wide pore (Benz and Hancock 1981) that allows passive diffusion of a broad range of small molecular weight hydrophilic materials (Nikaido 1979). In *E. coli* it is activated at high osmolarities of the surrounding medium (Forst *et al.* 1988). The high transcription during incubation in the light compared with the transcription in the dark where the epibiont is unable to perform photosynthesis leads to the conclusion that the OmpC protein is promoting the uptake of excreted substrates from the epibiont.

### 5.2.3 Dependence of the central bacterium on the epibiont

The genome analysis revealed the central bacterium to be a heterotrophic bacterium. It lacks anaerobic respiration and fermentation pathways and therefore uses glycolysis, the citrate cycle and a respiration chain to generate ATP. This however requires oxygen as terminal electron acceptor which is not present under the anoxic conditions used for the cultivation of the phototrophic consortia enrichment culture. In that regard, it is interesting to find a cytochrome *bd* quinol oxidase transcribed in the central bacterium. The *Escherichia coli* cytochrome *bd* oxidase has a very high affinity for oxygen (D'Mello *et al.* 1996). The enzyme

is a heterodimer encoded by *cydA* and *cydB* (Miller *et al.* 1988) and appears to be optimally expressed under anaerobic or near anaerobic conditions (Brondsted and Atlung 1996). Although the cultivation of the phototrophic enrichment culture is taking place under anoxic conditions, it cannot be ruled out that minute amounts of oxygen are present in the medium. These conditions resemble the chemocline, the natural habitat of the phototrophic consortium, in which oxygen is available in the upper and sulfide in the lower water layers (Overmann *et al.* 1998b). In this environment, with an oxidase highly affine for oxygen, even low concentrations of oxygen can be used by the central bacterium as terminal electron acceptor.

The fact that the central bacterium possesses all characteristics of an autonomous heterotrophic bacterium raises the question why it becomes almost inactive in the dark although substrates that could be metabolized are present. As seen in transcriptome analysis, the mRNA content dropped to a fraction of the mRNA expressed in the light and incubation of phototrophic consortia in the dark with  $^{13}\text{C}$ -labeled 2-oxoglutarate or leucine revealed no incorporation of labeled carbon into the central bacterium. This effect has also been observed in consortia of chemocline water samples using  $[\text{U-}^{14}\text{C}]$  2-oxoglutarate as substrate (Glaeser and Overmann 2003a). Only in the presence of both light and sulfide incorporation of  $^{14}\text{C}$  was measured. If either light or sulfide was missing, incorporation into phototrophic consortia dropped from 87 % to less than 1.4 %. This inactivity during periods of darkness might point towards a high synchronization of the symbiotic partner's metabolisms. Since the epibionts and the central bacterium of "*C. aggregatum*" divide together coordinately to form two complete consortia (Overmann and Schubert 2002), being metabolically highly active would be of no advantage to the central bacterium since cell division is only possible in unison with the epibiont. The synchronization of their metabolisms shows the high coordination in this interspecies interaction. A similar case of self regulation was observed in the bacteria-bacteria symbiosis between the thermophilic bacterium *Symbiobacterium thermophilum* that lives in a syntrophic association with a *Bacillus* strain (Ueda and Beppu 2007). *S. thermophilum* lacks the carbonic anhydrase that provides bicarbonate for enzymatic carboxylation reactions and thus depends on high  $\text{CO}_2$  concentrations generated by the *Bacillus* symbiont (Ueda *et al.* 2004; Watsuji *et al.* 2006). In the absence of the symbiont, *S. thermophilum* produces alkaloid self-growth inhibitors (Watsuji *et al.* 2006; Watsuji *et al.* 2007) that are removed by the *Bacillus* strain reinitiating growth in the partner organism.

### 5.2.4 Evidence of metabolite transfer from the central bacterium to the epibiont

With the results presented in this work, advantages resulting for the central bacterium of the close association with a green sulfur bacterium have initially been demonstrated. In the case of the epibiont, advantages known to date gained from the symbiosis with its heterotrophic partner, derive solely from the motility of the central bacterium.

Since green sulfur bacteria are not growing strictly autotrophically, but can incorporate carbon substrates like acetate and propionate in the presence of CO<sub>2</sub> (Overmann 2000) for mixotrophic growth, the availability of those substrates provide an advantage to GSB. For the epibiont it is known that it can metabolize acetate that facilitates growth during cultivation in pure culture (Vogl *et al.* 2006). According to transcriptome analysis, a pyruvate-ferredoxin oxidoreductase, a phosphotransacetylase, an acetate kinase and a hydrogenase are transcribed in the central bacterium. With this set of enzymes, pyruvate can be metabolized to acetate via acetyl-CoA and acetylphosphate. Thereby, one ATP and one NADH could be gained by the central bacterium through oxidative phosphorylation and acetate, CO<sub>2</sub> and H<sub>2</sub> would be gained, of which the first two are substrates of the epibiont (Vogl *et al.* 2006). This hypothetical pathway is in line with the finding that <sup>13</sup>C from externally added [U-<sup>13</sup>C] 2-oxoglutarate is also found in the epibiont after 3 h of incubation in the light. In this way, the enhanced incorporation of H<sup>14</sup>CO<sub>3</sub><sup>-</sup> into the associated epibiont in the presence of the amino acids (Fig. 10. a, b) alanine, arginine, histidine, methionine, proline, serine and valine in comparison with the positive control might be explained by an increased rate of metabolism. According to analyses of carbon flux through *Chl. tepidum* (Tang and Blankenship 2010), acetate can be assimilated via both of the reverse TCA as well as the oxidative TCA cycle. The oxidative TCA cycle revealed herein may explain better cell growth during mixotrophic growth with acetate, as energy is generated through the oxidative TCA cycle.

To date, it is not known why green sulfur bacteria excrete carbon substrates, but the association with a heterotrophic bacterium might enable the epibiont to cycle these substrates in order to retrieve acetate in return, facilitating its growth.

### 5.2.5 Implications for the origin of phototrophic consortia

Looking at the complex interaction between the two symbiotic partners of "*C. aggregatum*", the advantages gained by both species make it easy to understand why they form an evolutionary stable interaction. However, the advent of this interspecies interaction is not well

understood, since the benefits for both partners derive directly or indirectly from the directed movement towards light and sulfide. The orientation of the consortium towards light and sulfide is of special interest since probably neither of these attractants is used in the metabolism of the motile central bacterium. From the perspective of the epibiont, relying on a non-photosynthetic partner for transportation would pose the risk of being carried into dark, and/or sulfide-free unfavorable deeper water layers. Therefore, for the association to be advantageous for both organisms, the ability to sense these external stimuli had to be already present in the central bacterium at the origin of consortia formation. As our data have shown for the epibiont and as it has been documented for GSB (Sirevag and Ormerod 1970), substantial amounts of carbon are excreted by these organisms. Since GSB usually form layers in the chemocline of stratified water bodies (Overmann 2000), sensing light, oxygen and sulfide would enable heterotrophic bacteria to orientate themselves towards the habitat of GSB where carbon substrates are available. Phototrophic consortia have shown to react to all three of these stimuli (light, sulfide (Frösl and Overmann 1998; Glaeser and Overmann 2003b, 2003a), oxygen (Müller and Overmann, unpublished data). With the discovery of 32 PAS domain proteins in the central bacterium, a variety of parameters including small molecules, ions, gases, light and redox state (Henry and Crosson 2011) could be detected. Transcriptome analysis revealed that with the exception of the bacteriophytochrome, all of these genes are at least transcribed at a basic level. Furthermore, of the few proteins that were expressed higher during incubation in the dark, four proteins were involved in signal transduction, one being a PAS domain protein. In the epibiont, only 3 PAS domains are encoded in the genome. In stratified lakes, gradients of light and oxygen or sulfide are nearly omnipresent, whereas gradients of organic matter usually have a very limited radius (Krems *et al.* 1998). In the same manner, magnetotactic bacteria use the earth magnetic field to orientate along geomagnetic field lines by the permanent magnetic dipole moment imparted by magnetosomes. Together with chemotaxis, aerotaxis, and perhaps phototaxis, they locate and maintain their growth-favoring position in vertical chemical gradients within their aquatic habitats (Jogler and Schüler 2009).

In addition to the PAS domains, recent ultrastructural studies of the central bacterium of "*C. aggregatum*" revealed a conspicuous 35-nm-thick and up to 1  $\mu$ m-long zipper-like crystalline structures were found that resemble the chemotaxis receptor Tsr of *Escherichia coli* (Wanner *et al.* 2008). In a comparative ultrastructural study of 13 distantly related organisms harboring chemoreceptor arrays from all 7 major signaling domain classes,

receptors were found to possess an universal structure which has presumably been conserved over long evolutionary distances (Briegel et al. 2009). The prominent ultrastructure discovered in the central bacterium exhibits several similarities to the chemoreceptors reported. It therefore provides a first indication that chemotaxis genes are not only present in the central bacterial genome but that a chemotaxis receptor is expressed in the central bacterium.

With the initial ability of the central bacterium to sense light, sulfide and oxygen, the essential building block for the coevolution of the two organisms would be present at the beginning of consortia formation, enabling the development towards a highly complex mutualistic interspecies interaction.

### 5.2.6 Features of the epibiont that relate to symbiosis

In order to make a first assessment of the imprint of symbiotic lifestyle on the genome of the epibiont, the genome features of *Chl. chlorochromatii* CaD can be compared to those of the archaeal interspecies consortium of *Nanoarchaeum equitans* and *Ignicoccus hospitalis*.

*N. equitans* is the representative of a new archaeal phylum *Nanoarchaeota* and most likely represents a parasitic epibiont of *Ignicoccus*. The *N. equitans* genome comprises only 491 kb and encodes 552 genes, rendering it the smallest genome for an exosymbiont known to date (Waters et al. 2003). Genome reduction includes almost all genes required for the *de novo* biosynthesis of amino acids, nucleotides, cofactors, and lipids as well as many known pathways for carbon assimilation. Commensurate with this findings, the *Nanoarchaeum equitans* epibiont appears to acquire its lipids from its host (Waters et al. 2003). Yet, the *N. equitans* genome harbors only few pseudogenes or regions of non-coding DNA compared with genomes of obligate bacterial symbionts that undergo reductive evolution, and thus is genomically significantly more stable than other obligate parasites. This has been interpreted as evidence for a very ancient relationship between *N. equitans* and *Ignicoccus* (Brochier et al. 2005). The recent bioinformatic analysis also suggest that *N. equitans* represents a derived, fast-evolving euryarchaeal lineage rather than the representative of a deep-branching basal phylum (Brochier et al. 2005). With a size of 1.30 Mbp and 1,494 predicted ORFs, the genome of *I. hospitalis* shows a pronounced genome reduction that has been attributed to the reduced metabolic complexity of its anaerobic and autotrophic lifestyle and an highly efficient adaptation to the low energy yield of its metabolism (Podar et al. 2008).

Similar to *I. hospitalis*, the genome size of *Pelagibacter ubique* (1.31 Mbp and 1,354 ORFs) so far marks the lower limit of free-living organisms. It exhibits hallmarks of genome streamlining such as the lack of pseudogenes, introns, transposons, extrachromosomal elements, only few paralogs and the shortest intergenic spacers yet observed for any cell (Giovannoni et al. 2005). This is likely to reduce the costs of cellular replication (Mira et al. 2001). As a second feature, the genome of *P. ubique* has a G:C-content of 29.7% that may decrease its cellular requirements for fixed nitrogen (Dufresne et al. 2005).

By comparison, a much larger genome size of 2.57 Mbp has been determined for *Chl. chlorochromatii* CaD. This size represents the average value of the 11 other publicly available genomes of green sulfur bacteria. Thus, a reduction in genome size that is characteristic for bacterial endosymbionts (Andersson and Kurland 1998; Moran et al. 2002; Moran et al. 2008) and for the archaeal consortium (Podar et al. 2008) did not occur during the evolution of the epibiont of "*C. aggregatum*". This suggests (i) a shorter period of co-evolution of the two partner bacteria in phototrophic consortia, (ii) a significantly slower rate of evolution of their genomes or (iii) that the genome of *Chlorobium chlorochromatii* is not undergoing a streamlining process as observed in other symbiotic associations. The latter suggestion would indicate a lack of selective advantage for "*Chl. aggregatum*" from genome streamlining.

Wet-lab and *in silico* analyses of the epibiont genome revealed the presence of several putative symbiosis genes in the epibiont. An initial combination of suppression subtractive hybridization with bioinformatics approaches identified four open reading frames (ORFs) as candidates. Two of the ORFs (Cag0614 and 0616) exhibit similarities to putative filamentous haemagglutinins that harbor RGD (arginine-glycine-aspartate) tripeptides. In pathogenic bacteria, haemagglutinins with these motifs are involved in the attachment to mammalian cells (Vogl et al. 2008). Most notably, a comparative study of 580 sequenced prokaryotic genomes revealed that Cag 0614 and 0616 represent the largest genes detected in prokaryotes so far. In fact, Cag 0616 is only surpassed in length by the exons of the human titin gene (Reva and Tümmler 2008). The two other genes detected (Cag1919 and 1920) resembles RTX (repeats in toxin) -like proteins and haemolysins, respectively. All four genes have in common that they are unique in *Chl. chlorochromatii* CaD and that certain domains of their inferred products are only known from bacterial virulence factors. If the four genes were not misassigned, they are potentially involved in the symbiotic interaction between the two partner bacteria in phototrophic consortia.

To identify additional symbiosis genes, *in silico* subtractive hybridization between the genome sequence of *Chl. chlorochromatii* CaD and the other 11 sequences of green sulfur bacterial genomes was performed. This yielded 186 ORFs unique to the epibiont, 99 of which encode for hypothetical proteins of yet unknown function. Although this provides a large number of putative symbiosis genes, the numbers are rather low compared to the unique and unknown ORFs in the other green sulfur bacteria (Table 8). Even if it is assumed that all of these unknown genes encode for proteins involved in symbiosis, this number is still rather small compared to the 1,387 genes encoding niche-specific functions in enterohaemorrhagic *E. coli* O157:H7 (Perna et al 2001). Low numbers of niche-specific genes have been reported for *Salmonella enterica* or *Bacillus anthracis* and have been interpreted as indication for preadaptation of the non-pathogenic ancestor. Indeed, several of the physiological characteristics of green sulfur bacteria are regarded as preadaptive traits for interactions with other prokaryotes.

One feature of green sulfur bacteria which provides interaction with other prokaryotes is their carbon metabolism. Green sulfur bacteria autotrophically assimilate CO<sub>2</sub> through the reductive tricarboxylic acid cycle. One instantaneous product of photosynthetic fixation of CO<sub>2</sub> is 2-oxoglutarate, and 2-oxo acids represent typical excretion products of photosynthesizing cells (Sirevag and Ormerod 1970). In natural environments, *Chlorobium limicola* excretes significant amounts of photosynthetically fixed carbon (Czeczuga and Gradzki 1973) and thus constitutes a potential electron donor for associated bacteria. Excretion of organic carbon compounds has also been demonstrated for *Chlorobium chlorochromatii* strain CaD (Fig. 16 and Pfannes (2007)). *Vice versa*, green sulfur bacteria can also take advantage of organic carbon compounds produced by other, e.g. fermenting, bacteria. During phototrophic growth, they are capable of assimilating pyruvate as well as acetate and propionate through reductive carboxylation in the presence of CO<sub>2</sub> (pyruvate:ferredoxin oxidoreductase; Uyeda and Rabinowitz 1971) or HCO<sub>3</sub><sup>-</sup> (phosphoenolpyruvate carboxylase; Chollet et al. 1996). The assimilation of organic carbon compounds reduces the amount of electrons required per unit cellular carbon synthesized. This capability thus enhances photosynthetic growth yield and results in a competitive advantage for green sulfur bacteria.

In their natural environment, green sulfur bacteria are limited to habitats where light reaches anoxic bottom waters such as in thermally stratified or meromictic lakes. Here, cells encounter conditions favorable for growth exclusively in a rather narrow (typically cm to dm



thick) zone of overlap between light and sulfide. Compared to other phototrophs, green sulfur bacteria are extremely low-light adapted and capable of exploiting minute light quantum fluxes by their extraordinarily large photosynthetic antenna complexes, the chlorosomes. Until recently, the heterogeneity of pigments complicated the identification of the structural composition of chlorosomes. When a *Chlorobaculum tepidum* triple mutant that almost exclusively harbors BChl *d* was constructed, a *syn-anti* stacking of monomers and self-assembly of bacteriochlorophylls into tubular elements could be demonstrated within the chlorosomes (Ganapathy et al. 2009).

Since they minimize the energetically costly protein synthesis, chlorosomes represent the most effective light harvesting system known. Up to  $215,000 \pm 80,000$  bacteriochlorophyll molecules (in *Chlorobaculum tepidum*; Montano et al. 2003) can constitute a single chlorosome, that is anchored to five to ten reaction centers in the cytoplasmic membrane (Amesz 1991). This ratio of chlorophyll to reaction center is orders of magnitudes higher compared with other photosynthetic antenna structures. In the phycobilisomes of cyanobacteria the ratio is 220:1 (Clement-Metral et al. 1985), 100-140:1 in light harvesting complex II of anoxygenic phototrophic proteobacteria (Van Grondelle et al. 1983) and 28:1 in the light harvesting complex I (Melkozernov et al. 2006). The enormous size of the photosynthetic antenna of green sulfur bacteria enables them to colonize extreme low-light habitats up to depths of 100 m in the Black Sea (Marschall et al. 2010; Overmann et al. 1992) or below layers of other phototrophic organisms like purple sulfur bacteria (Pfennig 1978). Commensurate with their adaptation to extreme light limitation, green sulfur bacteria also exhibit a significantly reduced maintenance energy requirement compared to other bacteria (Overmann et al. 1992; Veldhuis and van Gemerden 1986). *Chlorobium* phylotype BS-1 isolated from the Black Sea maintained a constant level of cellular ATP over 52 days, if exposed to low light intensities of  $0.01 \text{ mmol quanta m}^{-2} \text{ s}^{-1}$  (Marschall et al. 2010). Their effective antenna system and reduced maintenance energy requirement allows green sulfur bacteria to colonize habitats in which other photosynthetic bacteria are unable to grow. This flexibility might have been of advantage for the green sulfur bacteria during consortia formation, since the symbiosis with the heterotrophic partner organism comprised motility and thereby the possibility to move between different layers of water.

### 5.3 Conclusions

While different types of symbioses and syntrophic associations all provide an energetic advantage to one or both partners, only few of these associations reached the level of organizational complexity of the highly-structured, permanent consortia. Thus, a permanent cell-cell contact is not mandatory in the case of syntrophic cultures in which depletion of substrates can lead to disaggregation of the associations (Peduzzi *et al.* 2003). The highly developed and obligate interaction in phototrophic consortia is likely to be related to the pronounced energy limitation in the low-light habitats and to the efficient and regulated exchange of metabolites. Our extensive studies illuminating the interaction of the central bacterium and the epibiont of "*C. aggregatum*" with a focused attention towards the exchange of metabolites, painted a complex picture of substrate utilization within the phototrophic consortium. Therein, substrates provided by the epibiont competitively compete for the uptake by the central bacterium with external substrates in the surrounding medium. This implies that uptake of substrates is mediated by the multitude of transcribed transporters identified in the central bacterium and not by the periplasmic tubuli spanning the interspace between central bacterium and epibiont (Wanner *et al.* 2008). In this regard, a future assignment in consortia research should include the localization of transporters in both the central bacterium and the epibiont, in which internal sorting renders the contact site to the central bacterium void of chlorosomes (Vogl *et al.* 2006). Therefore, this area might be involved in the rapid exchange of metabolites observed in the phototrophic consortium. Amino acids have been shown to be one group of substrates exchanged between the symbiotic partners, however, several lines of evidence give rise to the conclusion that a whole variety of substrates is exchanged. Thereby, rapid cell-to-cell signaling could be achieved within the symbiosis, explaining the high coordination between the two partners that is necessary for a coordinated division or metabolic coupling.

As it has been discussed, GSB possess several traits of preadaptation to a symbiotic lifestyle. It is therefore of importance to the understanding of the advent of phototrophic consortia to investigate the origin and radiation of this family. The accumulation and analysis of the available sequence data on GSB revealed, that the majority of the diversity is already covered, although isolates of certain clusters are still missing. Further research should therefore concentrate more on the optimization of the current culturing techniques than on searching for so far unfound sequence types. In that regard, it would be of importance to

isolate epibionts from different phototrophic consortia to understand the genetic adaptations necessary for green sulfur bacteria to become an epibiont. As the example of green or brown pigmentation has shown, clusters of genes that impact the adaptation of GSB to their habitat, are transferred between green sulfur bacterial strains. Since the gain of motility by the epibiont seems to constitute a selective advantage, clusters of genes that promote the attachment of GSB to a motile betaproteobacterium might initiate the co-evolution between these two different species.

## References

- Alexander, B., Andersen, J. H., Cox, R. P., and Imhoff, J. F. (2002). Phylogeny of green sulfur bacteria on the basis of gene sequences of 16S rRNA and of the Fenna-Matthews-Olson protein. *Arch Microbiol* **178**(2), 131-40.
- Amesz, J. (1991). Green photosynthetic bacteria and heliobacteria. *Variations in Autotrophic Life*, 99-119.
- Andersson, S. G. E., and Kurland, C. G. (1998). Ancient and recent horizontal transfer events: The origins of mitochondria. *APMIS, Supplement* **106**(84), 5-14.
- Anne Isabel, B. (2007). Neue Ansätze zur Analyse der bakteriellen Interaktionen in phototrophen Konsortien. Diploma thesis at the Ludwig-Maximilians Universität München; 180p
- Benz, R., and Hancock, R. E. (1981). Properties of the large ion-permeable pores formed from protein F of *Pseudomonas aeruginosa* in lipid bilayer membranes. *Biochim Biophys Acta* **646**(2), 298-308.
- Biebl, H., and Pfennig, N. (1978). Growth yields of green sulfur bacteria in mixed cultures with sulfur and sulfate reducing bacteria. *Archives of microbiology* **117**(1), 9-16.
- Briegel, A., Ortega, D. R., Tocheva, E. I., Wuichet, K., Zhuo, L., Songye, C., Müller, A., Iancu, C. V., Murphy, G. E., Dobro, M. J., Zhulin, I. B., and Jensen, G. J. (2009). Universal architecture of bacterial chemoreceptor arrays. *Proceedings of the National Academy of Sciences of the United States of America* **106**(40), 17181-17186.
- Brochier, C., Gribaldo, S., Zivanovic, Y., Confalonieri, F., and Forterre, P. (2005). Nanoarchaea: representatives of a novel archaeal phylum or a fast-evolving euryarchaeal lineage related to Thermococcales? *Genome biology* **6**(5).
- Brockmann, H., Jr. (1976). Bacteriochlorophyll e: structure and stereochemistry of a new type of chlorophyll from *Chlorobiaceae*. *Philosophical transactions of the Royal Society of London. Series B: Biological sciences* **273**, 277-85.
- Brondsted, L., and Atlung, T. (1996). Effect of growth conditions on expression of the acid phosphatase (cyx-appA) operon and the appY gene, which encodes a transcriptional activator of *Escherichia coli*. *J Bacteriol* **178**(6), 1556-64.
- Caldwell, D. E., and Tiedje, J. M. (1975). The structure of anaerobic bacterial communities in the hypolimnia of several Michigan lakes. *Canadian Journal of Microbiology* **21**(3), 377-385.

- Castenholz, R. W. (1978). The biogeography of hot spring algae through enrichment cultures. *Mitt.Int.Ver.Limnol.*(21), 296-315.
- Chollet, R., Vidal, J., and O'Leary, M. H. (1996). Phosphoenolpyruvate carboxylase: A ubiquitous, highly regulated enzyme in plants. *Annual Review of Plant Physiology and Plant Molecular Biology* **47**(1), 273-298.
- Clement-Metral, J. D., Gantt, E., and Redlinger, T. (1985). A photosystem II-phycobilisome preparation from the red alga, *Porphyridium cruentum* : Oxygen evolution, ultrastructure, and polypeptide resolution. *Archives of Biochemistry and Biophysics* **238**(1), 10-17.
- Coffin, R. B. (1989). Bacterial uptake of dissolved free and combined amino acids in estuarine waters. *limnol Oceanogr* **34**(3), 531-542.
- Czeczuga, B., and Gradzki, F. (1973). Relationship between extracellular and cellular production in the sulphuric green bacterium *Chlorobium limicola* Nads. (Chlorobacteriaceae) as compared to primary production of phytoplankton. *Hydrobiologia* **42**(1), 85-95.
- D'Mello, R., Hill, S., and Poole, R. K. (1996). The cytochrome bd quinol oxidase in *Escherichia coli* has an extremely high oxygen affinity and two oxygen-binding haems: implications for regulation of activity in vivo by oxygen inhibition. *Microbiology* **142** ( Pt 4), 755-63.
- Davenport, C., Ussery, D., and Tümmler, B. (2010). Comparative genomics of green sulfur bacteria. *Photosynthesis Research* **104**(2), 137-152.
- de Castro, A., Araújo, S., Reis, A., Moura, R., Francini-Filho, R., Pappas, G., Rodrigues, T., Thompson, F., and Krüger, R. (2010). Bacterial Community Associated with Healthy and Diseased Reef Coral <i>Scleractinia</i> from Eastern Brazil. *Microbial Ecology* **59**(4), 658-667.
- Dufresne, A., Garczarek, L., and Partensky, F. (2005). Accelerated evolution associated with genome reduction in a free-living prokaryote. *Genome biology* **6**(2).
- Forst, S., Delgado, J., Ramakrishnan, G., and Inouye, M. (1988). Regulation of ompC and ompF expression in *Escherichia coli* in the absence of envZ. *J Bacteriol* **170**(11), 5080-5.
- Frias-Lopez, J., Zerkle, A. L., Bonheyo, G. T., and Fouke, B. W. (2002). Partitioning of Bacterial Communities between Seawater and Healthy, Black Band Diseased, and Dead Coral Surfaces. *Applied and environmental microbiology* **68**(5), 2214-2228.
- Fröstl, J. M., and Overmann, J. (1998). Physiology and tactic response of the phototrophic consortium "*Chlorochromatium aggregatum*". *Arch Microbiol* **169**(2), 129-35.

- Fröstl, J. M., and Overmann, J. (2000). Phylogenetic affiliation of the bacteria that constitute phototrophic consortia. *Arch Microbiol* **174**(1-2), 50-8.
- Ganapathy, S., Sengupta, S., Wawrzyniak, P. K., Huber, V., Buda, F., Baumeister, U., Würthner, F., and de Groot, H. J. M. (2009). Zinc chlorins for artificial light-harvesting self-assemble into antiparallel stacks forming a microcrystalline solid-state material. *Proceedings of the National Academy of Sciences* **106**(28), 11472-11477.
- Giovannoni, S. J., Tripp, H. J., Givan, S., Podar, M., Vergin, K. L., Baptista, D., Bibbs, L., Eads, J., Richardson, T. H., Noordewier, M., Rapp, M. S., Short, J. M., Carrington, J. C., and Mathur, E. J. (2005). Genetics: Genome streamlining in a cosmopolitan oceanic bacterium. *Science* **309**(5738), 1242-1245.
- Glaeser, J., and Overmann, J. (2003a). Characterization and in situ carbon metabolism of phototrophic consortia. *Applied and environmental microbiology* **69**(7), 3739-3750.
- Glaeser, J., and Overmann, J. (2003b). The significance of organic carbon compounds for in situ metabolism and chemotaxis of phototrophic consortia. *Environmental microbiology* **5**(11), 1053-1063.
- Glöckner, F. O., Zaichikov, E., Belkova, N., Denissova, L., Pernthaler, J., Pernthaler, A., and Amann, R. (2000). Comparative 16S rRNA Analysis of Lake Bacterioplankton Reveals Globally Distributed Phylogenetic Clusters Including an Abundant Group of Actinobacteria. *Applied and environmental microbiology* **66**(11), 5053-5065.
- Gorlenko, V. M., and Kuznetsov, S. I. (1971). Vertical distribution of photosynthesizing bacteria in lake Konon"er in Mari ASSR. *Mikrobiologiya* **40**(4), 746-747.
- Harrison, M. J. (2005). Signaling in the arbuscular mycorrhizal symbiosis. *Annu Rev Microbiol* **59**, 19-42.
- Hedlund, B. P., and Staley, J. T., Eds. (2003). Microbial endemism and biogeography. Microbial Diversity and Bioprospecting. Edited by A. T. Bull. Washington DC.
- Henry, J. T., and Crosson, S. (2011). Ligand-binding PAS domains in a genomic, cellular, and structural context. *Annu Rev Microbiol* **65**, 261-86.
- Imhoff, J. F. (2001). True marine and halophilic anoxygenic phototrophic bacteria. *Arch Microbiol* **176**(4), 243-54.
- Imhoff, J. F. (2003). Phylogenetic taxonomy of the family Chlorobiaceae on the basis of 16S rRNA and fmo (Fenna-Matthews-Olson protein) gene sequences. *International Journal of Systematic and Evolutionary Microbiology* **53**(4), 941-951.

- Imhoff, J. F., and Thiel, V. (2010). Phylogeny and taxonomy of Chlorobiaceae. *Photosynth Res* **104**(2-3), 123-36.
- Jahn, U., Gallenberger, M., Paper, W., Junglas, B., Eisenreich, W., Stetter, K. O., Rachel, R., and Huber, H. (2008). Nanoarchaeum equitans and Ignicoccus hospitalis: new insights into a unique, intimate association of two archaea. *J Bacteriol* **190**(5), 1743-50.
- Jogler, C., and Schuler, D. (2009). Genomics, genetics, and cell biology of magnetosome formation. *Annu Rev Microbiol* **63**, 501-21.
- Kiers, E. T., Duhamel, M., Beesetty, Y., Mensah, J. A., Franken, O., Verbruggen, E., Fellbaum, C. R., Kowalchuk, G. A., Hart, M. M., Bago, A., Palmer, T. M., West, S. A., Vandenkoornhuyse, P., Jansa, J., and Bucking, H. (2011). Reciprocal rewards stabilize cooperation in the mycorrhizal symbiosis. *Science* **333**(6044), 880-2.
- Kirk, P. M., Cannon, P. F., David, J. C., and Stalpers, J. (2001). "Ainsworth and Bisby's Dictionary of the Fungi." (9th ed.) CAB International, Wallingford, UK.
- Kosten, S., Huszar, V. L., Mazzeo, N., Scheffer, M., Sternberg Lda, S., and Jeppesen, E. (2009). Lake and watershed characteristics rather than climate influence nutrient limitation in shallow lakes. *Ecol Appl* **19**(7), 1791-804.
- Krembs, C., Juhl, A. R., Long, R. A., and Azam, F. (1998). Nanoscale patchiness of bacteria in lake water studied with the spatial information preservation method. *Limnology and Oceanography* **43**(2), 307-314.
- Kroer, N., Jorgensen, N. O., and Coffin, R. B. (1994). Utilization of dissolved nitrogen by heterotrophic bacterioplankton: a comparison of three ecosystems. *Appl Environ Microbiol* **60**(11), 4116-23.
- Larson, C. R., Seng, C. O., Lauman, L., Matthies, H. J., Wen, J., Blankenship, R. E., and Allen, J. P. (2011). The three-dimensional structure of the FMO protein from Pelodictyon phaeum and the implications for energy transfer. *Photosynth Res* **107**(2), 139-50.
- Lauterborn, R. (1906). Zur Kenntnis der sapropelischen Flora. *Allg.Bot.Z.*(12), 196-197.
- Lodwig, E. M., Hosie, A. H., Bourdes, A., Findlay, K., Allaway, D., Karunakaran, R., Downie, J. A., and Poole, P. S. (2003). Amino-acid cycling drives nitrogen fixation in the legume-Rhizobium symbiosis. *Nature* **422**(6933), 722-6.
- Maresca, J. A., Gomez Maqueo Chew, A., Ponsat+-, M. R., Frigaard, N. U., Ormerod, J. G., and Bryant, D. A. (2004). The bchU Gene of Chlorobium tepidum Encodes the C-20 Methyltransferase in Bacteriochlorophyll c Biosynthesis. *Journal of bacteriology* **186**(9), 2558-2566.

- Marschall, E., Jogler, M., Henssge, U., and Overmann, J. (2010). Large-scale distribution and activity patterns of an extremely low-light-adapted population of green sulfur bacteria in the Black Sea. *Environmental microbiology* **12**(5), 1348-1362.
- Massana, R., DeLong, E. F., and Pedrós-Alió, C. (2000). A Few Cosmopolitan Phylotypes Dominate Planktonic Archaeal Assemblages in Widely Different Oceanic Provinces. *Applied and environmental microbiology* **66**(5), 1777-1787.
- May, L., Spears, B. M., Dudley, B. J., and Hatton-Ellis, T. W. (2010). The importance of nitrogen limitation in the restoration of Llangorse Lake, Wales, UK. *J Environ Monit* **12**(1), 338-46.
- Melkozernov, A. N., Barber, J., and Blankenship, R. E. (2006). Light harvesting in photosystem I supercomplexes. *Biochemistry* **45**(2), 331-345.
- Migal, M. G. (2011). The cladoceran trophic status in the nitrogen limited ecosystem of Lake Kinneret (Israel). *J Environ Biol* **32**(4), 455-62.
- Miller, M. J., Hermodson, M., and Gennis, R. B. (1988). The active form of the cytochrome d terminal oxidase complex of *Escherichia coli* is a heterodimer containing one copy of each of the two subunits. *J Biol Chem* **263**(11), 5235-40.
- Mira, A., Ochman, H., and Moran, N. A. (2001). Deletional bias and the evolution of bacterial genomes. *Trends in Genetics* **17**(10), 589-596.
- Montano, G. A., Bowen, B. P., LaBelle, J. T., Woodbury, N. W., Pizziconi, V. B., and Blankenship, R. E. (2003). Characterization of *Chlorobium tepidum* chlorosomes: A calculation of bacteriochlorophyll c per chlorosome and oligomer modeling. *Biophysical Journal* **85**(4), 2560-2565.
- Montesinos, E., Guerrero, R., Abella, C., and Esteve, I. (1983). Ecology and Physiology of the Competition for Light Between *Chlorobium limicola* and *Chlorobium phaeobacteroides* in Natural Habitats. *Appl Environ Microbiol* **46**(5), 1007-16.
- Moran, N. A., McCutcheon, J. P., and Nakabachi, A. (2008). Genomics and evolution of heritable bacterial symbionts, Vol. 42, pp. 165-190.
- Moran, P. J., Cheng, Y., Cassell, J. L., and Thompson, G. A. (2002). Gene expression profiling of *Arabidopsis thaliana* in compatible plant-aphid interactions. *Archives of Insect Biochemistry and Physiology* **51**(4), 182-203.
- Mullineaux, C. W., Mariscal, V., Nenninger, A., Khanum, H., Herrero, A., Flores, E., and Adams, D. G. (2008). Mechanism of intercellular molecular exchange in heterocyst-forming cyanobacteria. *EMBO Journal* **27**(9), 1299-1308.



- Myers, J. L., and Richardson, L. L. (2009). Adaptation of cyanobacteria to the sulfide-rich microenvironment of black band disease of coral. *FEMS Microbiology Ecology* **67**(2), 242-251.
- Nikaido, H. (1979). Permeability of the outer membrane of bacteria. *Angew Chem Int Ed Engl* **18**(5), 337-50.
- Overmann, J. (2000). The family Chlorobiaceae. *The Prokaryotes. An Evolving Electronic Resource for the Microbiological Community, Vol Release 3.1*.
- Overmann, J. (2001). Green sulfur bacteria. *Bergey's Manual of Systematic Bacteriology* **1**, 601-605.
- Overmann, J., Cypionka, H., and Pfennig, N. (1992). An extremely low-light-adapted phototrophic sulfur bacterium from the Black Sea. *Limnology & Oceanography* **37**(1), 150-155.
- Overmann, J., and Schubert, K. (2002). Phototrophic consortia: Model systems for symbiotic interrelations between prokaryotes. *Archives of microbiology* **177**(3), 201-208.
- Overmann, J., and Tuschak, C. (1997). Phylogeny and molecular fingerprinting of green sulfur bacteria. *Archives of microbiology* **167**(5), 302-309.
- Overmann, J., Tuschak, C., Sass, H., and Fröstl, J. (1998a). The ecological niche of the consortium "Pelochromatium roseum". *Archives of microbiology* **169**(2), 120-128.
- Overmann, J., Tuschak, C., Sass, H., and Fröstl, J. (1998b). The ecological niche of the consortium "Pelochromatium roseum". *Arch Microbiol* **169**(2), 120-8.
- Pantos, O., Cooney, R. P., Le Tissier, M. D. A., Barer, M. R., O'Donnell, A. G., and Bythell, J. C. (2003). The bacterial ecology of a plague-like disease affecting the Caribbean coral *Montastrea annularis*. *Environmental microbiology* **5**(5), 370-382.
- Papke, R. T., Ramsing, N. B., Bateson, M. M., and Ward, D. M. (2003). Geographical isolation in hot spring cyanobacteria. *Environ Microbiol* **5**(8), 650-9.
- Parkin, T. B., and Brock, T. D. (1980). The effects of light quality on the growth of phototrophic bacteria in lakes. *Archives of microbiology* **125**(1-2), 19-27.
- Peduzzi, S., Tonolla, M., and Hahn, D. (2003). Isolation and characterization of aggregate-forming sulfate-reducing and purple sulfur bacteria from the chemocline of meromictic Lake Cadagno, Switzerland. *FEMS Microbiology Ecology* **45**(1), 29-37.
- Perna, N. T., Plunkett, G., Burland, V., Mau, B., Glasner, J. D., Rose, D. J., Mayhew, G. F., Evans, P. S., Gregor, J., Kirkpatrick, H. A., Pèrfai, G., Hackett, J., Klink, S., Boutin,

- A., Shao, Y., Miller, L., Grotbeck, E. J., Davis, N. W., Lim, A., Dimalanta, E. T., Potamouisis, K. D., Apodaca, J., Anantharaman, T. S., Lin, J., Yen, G., Schwartz, D. C., Welch, R. A., and Blattner, F. R. (2001). Genome sequence of enterohaemorrhagic *Escherichia coli* O157:H7. *Nature* **409**(6819), 529-533.
- Pfannes, K. R. (2007). Dissertation University of Munich.
- Pfennig, N. (1978). General physiology and ecology of photosynthetic bacteria. *The Photosynthetic Bacteria*, 3-18.
- Pfennig, N. (1980). Syntrophic mixed cultures and symbiotic consortia with phototrophic bacteria: A review. *Anaerobes and Anaerobic Infections*, 127-131.
- Pfennig, N., and Trüper, H. G. (1989). Anoxygenic phototrophic bacteria. *Bergey's Manual of Systematic Bacteriology* **3**, 1635-1709.
- Podar, M., Anderson, I., Makarova, K. S., Elkins, J. G., Ivanova, N., Wall, M. A., Lykidis, A., Mavromatis, K., Sun, H., Hudson, M. E., Chen, W., Deciu, C., Hutchison, D., Eads, J. R., Anderson, A., Fernandes, F., Szeto, E., Lapidus, A., Kyrpides, N. C., Saier, M. H., Richardson, P. M., Rachel, R., Huber, H., Eisen, J. A., Koonin, E. V., Keller, M., and Stetter, K. O. (2008). A genomic analysis of the archaeal system *Ignicoccus hospitalis*-*Nanoarchaeum equitans* *Genome biology* **9**(11).
- Preiss, J. (1984). Bacterial glycogen synthesis and its regulation. *Annu Rev Microbiol* **38**, 419-58.
- Prell, J., White, J. P., Bourdes, A., Bunnewell, S., Bongaerts, R. J., and Poole, P. S. (2009). Legumes regulate *Rhizobium* bacteroid development and persistence by the supply of branched-chain amino acids. *Proceedings of the National Academy of Sciences of the United States of America* **106**(30), 12477-12482.
- Reva, O., and T•mmmler, B. (2008). Think big - Giant genes in bacteria. *Environmental microbiology* **10**(3), 768-777.
- Richardson, L. L., Miller, A. W., Broderick, E., Kaczmarzsky, L., Gantar, M., Stanic, D., and Sekar, R. (2009). Sulfide, microcystin, and the etiology of black band disease. *Dis Aquat Organ* **87**(1-2), 79-90.
- Sirevag, R., and Ormerod, J. G. (1970). Carbon dioxide fixation in green sulphur bacteria. *Biochemical Journal* **120**(2), 399-408.
- Sirevag, R., and Ormerod, J. G. (1977). Synthesis, storage and degradation of polyglucose in *Chlorobium thiosulfatophilum*. *Arch Microbiol* **111**(3), 239-44.

- Tang, K. H., and Blankenship, R. E. (2010). Both forward and reverse TCA cycles operate in green sulfur bacteria. *Journal of Biological Chemistry* **285**(46), 35848-35854.
- Triado-Margarit, X., Vila, X., and Abella, C. A. (2010). Novel green sulfur bacteria phylotypes detected in saline environments: ecophysiological characters versus phylogenetic taxonomy. *Antonie van Leeuwenhoek* **97**(4), 419-31.
- Tuschak, C., Glaeser, J., and Overmann, J. (1999). Specific detection of green sulfur bacteria by in situ hybridization with a fluorescently labeled oligonucleotide probe. *Archives of microbiology* **171**(4), 265-272.
- Tyler, B. (1978). Regulation of the assimilation of nitrogen compounds. *Annu Rev Biochem* **47**, 1127-62.
- Ueda, K., and Beppu, T. (2007). Lessons from studies of *Symbiobacterium thermophilum*, a unique syntrophic bacterium. *Biosci Biotechnol Biochem* **71**(5), 1115-21.
- Ueda, K., Yamashita, A., Ishikawa, J., Shimada, M., Watsuji, T. O., Morimura, K., Ikeda, H., Hattori, M., and Beppu, T. (2004). Genome sequence of *Symbiobacterium thermophilum*, an uncultivable bacterium that depends on microbial commensalism. *Nucleic Acids Res* **32**(16), 4937-44.
- Uyeda, K., and Rabinowitz, J. C. (1971). Pyruvate-ferredoxin oxidoreductase. IV. Studies on the reaction mechanism. *Journal of Biological Chemistry* **246**(10), 3120-3125.
- Van Grondelle, R., Hunter, C. N., Bakker, J. G. C., and Kramer, H. J. M. (1983). Size and structure of antenna complexes of photosynthetic bacteria as studied by singlet-singlet quenching of the bacteriochlorophyll fluorescence yield. *BBA - Bioenergetics* **723**(1), 30-36.
- Veldhuis, M. J. W., and van Gernerden, H. (1986). Competition between purple and brown phototrophic bacteria in stratified lakes: sulfide, acetate, and light as limiting factors. *FEMS Microbiology Letters* **38**(1), 31-38.
- Vogl, K., Glaeser, J., Pfannes, K. R., Wanner, G., and Overmann, J. (2006). *Chlorobium chlorochromatii* sp. nov., a symbiotic green sulfur bacterium isolated from the phototrophic consortium "Chlorochromatium aggregatum". *Arch Microbiol* **185**(5), 363-72.
- Wanner, G., Vogl, K., and Overmann, J. (2008). Ultrastructural characterization of the prokaryotic symbiosis in "*Chlorochromatium aggregatum*". *Journal of bacteriology* **190**(10), 3721-3730.

- Warthmann, R., Cypionka, H., and Pfennig, N. (1992). Photoproduction of H<sub>2</sub> from acetate by syntrophic cocultures of green sulfur bacteria and sulfur-reducing bacteria. *Archives of microbiology* **157**(4), 343-348.
- Waters, E., Hohn, M. J., Ahel, I., Graham, D. E., Adams, M. D., Barnstead, M., Beeson, K. Y., Bibbs, L., Bolanos, R., Keller, M., Kretz, K., Lin, X., Mathur, E., Ni, J., Podar, M., Richardson, T., Sutton, G. G., Simon, M., Soll, D., Stetter, K. O., Short, J. M., and Noordewier, M. (2003). The genome of *Nanoarchaeum equitans* : Insights into early archaeal evolution and derived parasitism. *Proceedings of the National Academy of Sciences of the United States of America* **100**(22), 12984-12988.
- Watsuji, T. O., Kato, T., Ueda, K., and Beppu, T. (2006). CO<sub>2</sub> supply induces the growth of *Symbiobacterium thermophilum*, a syntrophic bacterium. *Biosci Biotechnol Biochem* **70**(3), 753-6.
- Watsuji, T. O., Yamada, S., Yamabe, T., Watanabe, Y., Kato, T., Saito, T., Ueda, K., and Beppu, T. (2007). Identification of indole derivatives as self-growth inhibitors of *Symbiobacterium thermophilum*, a unique bacterium whose growth depends on coculture with a *Bacillus* sp. *Appl Environ Microbiol* **73**(19), 6159-65.
- Wenter, R., H • tz, K., Dibbern, D., Li, T., Reisinger, V., Ploscher, M., Eichacker, L., Eddie, B., Hanson, T., Bryant, D. A., and Overmann, J. (2010). Expression-based identification of genetic determinants of the bacterial symbiosis ' *Chlorochromatium aggregatum* '. *Environmental microbiology*.
- Wu, G. F., Wu, X. C., Xuan, X. D., and Zhou, X. P. (2006). Evaluation of nutrient limitation in aquatic ecosystems with nitrogen fixing bacteria. *J Environ Sci (China)* **18**(3), 537-42.



# I. Danksagung

An dieser Stelle möchte ich mich bei Herrn Prof. Dr. Jörg Overmann für die Bereitstellung dieses spannenden Themas und die gute wissenschaftliche Betreuung während der Doktorarbeit bedanken.

Für die gute Einweisung, die spannenden Diskussionen und die Entdeckung am Spaß des Konsortienprojekts möchte ich mich bei Kajetan Vogl, Roland Wenter und Kristina Pfannes ganz herzlich bedanken.

Ich danke der gesamten AG Overmann für die mir angebotene Hilfe in zahlreichen Projekten, die gute Zusammenarbeit und für die schöne Zeit die wir in, aber vor allem auch außerhalb der Arbeit miteinander verbracht haben. Besonders möchte ich mich auch für die lustigen Monate in der WG Overmann bei meinen damaligen Mitbewohnern Ovidiu Rücker und Mareike Jogler bedanken.

Mein herzlichster Dank gilt Katharina Hütz, mit der ich immer jegliches Leid und jegliche Freude über die Doktorarbeit teilen konnte und die mich während der gesamten Zeit motiviert hat.

Ganz besonders danke ich meinen Eltern die mich nicht nur während meiner Doktorarbeitszeit unterstützt haben, sondern bereits mein ganzes Leben für mich da waren.



## **III. Supplementary material**



**Suppl. Table 1.** References for species used in the phylogenetic study. `\*` sequences obtained in this study.

Species name	Strain	Strain collection no.	16S rDNA EMBL Accession no.	Length of sequence	Sampling site	16S-23S rRNA ITS	sigA	bchG	fmoA
<i>Cba. limnaeum</i>	UdG 6038	-	Y10641	1466	Wintergreen Lake, Michigan, USA (42.40 N 85.38 W)	AM049271*	-	AJ427485	-
<i>Cba. limnaeum</i>	UdG 6040	-	Y10642	1465	Font de La Puda, Spain (42.11 N 2.75 E)	AM049272*	AM050069*	AM050372*	*
<i>Cba. limnaeum</i>	UdG 6041	-	Y10643	1468	Lake Coromina, Spain (42.13 N 2.56 E)	AM049292*, AM049293*	AM050071*	AM050353*	AJ306183
<i>Cba. limnaeum</i>	UdG 6042	-	Y10644	1467	Lake Vilar, Spain (42.08 N 2.44 E)	AM049273*	AM050096*	AM050354*	-
<i>Cba. limnaeum</i>	UdG 6045	-	Y10646	1469	Wintergreen Lake, Michigan, USA (42.40 N 85.38 W)	AM049274*	AM050070*	AM050374*	*
<i>Cba. limnaeum</i>	1549	-	AJ299413	1387	Pond, Solling, Germany (N 51.72 E 9.62)	-	-	-	AJ306184
<i>Cba. limnaeum</i>	-	DSM 1677 <sup>T</sup>	AJ290831	1241	Lake Kinneret, Israel (32.83 N 35.58 E)	-	-	-	AJ391149
<i>Cba. limnaeum</i>	X	-	EF560700	1363	-	-	-	-	-
<i>Cba. limnaeum</i>	C	-	EF560701	1345	-	-	-	-	-
<i>Cba. macestae</i>	M	-	EF560696	1395	sulfidic spring, Black Sea Coast, Caucasus, Georgia (42.11 N 41.67 E)	-	-	-	*
<i>Cba. parvum</i>	8GSB	-	AJ888466	1298	salt pan, Gokarna, India (14.55 N	-	-	-	-

					74.31 E)				
<i>Cba. parvum</i>	NCIB 8327, UdG 6036	DSM 263 <sup>T</sup>	M31769/ Cpar_R0010	1504/ 1495	Garabogazkö, Turkmenistan (41.35 N 53.59 E)	AM049275*, AM049276*	AM050076*/C par_0407	AJ427487/C par_0576	AJ391147/ Cpar_1586
<i>Cba. parvum</i>	Ch38	-	AM940937	687	Orissa, Chilka lagoon, India (19.72 N 85.33 E)	-	-	-	-
<i>Cba. parvum</i>	UdG 6501Lms	-	EF064316	1270	Massona Lagoon (Parc Natural dels Aiguamolls de l'Empordà), Spain (42.21 N 3.10 E)	-	-	-	-
<i>Cba. parvum</i>	5Sg	-	AM690802	1270	marine tidal water, Konark, India (19.89 N 86.09 E)	-	-	-	-
<i>Cba. parvum</i>	NCIB 8346	-	AJ290830	1243	-	-	-	-	AJ391161
<i>Cba. parvum</i>	L	-	EF560699	1306	-	-	-	-	-
<i>Cba. tepidum</i>	TLS	ATCC 49652 <sup>T</sup>	M58468/CTrr naB16S	1453/ 1449	Travelodge, Rotorua, New Zealand (38.14 S 176.25 E)	NC_002932, 140937-141461	U67718	AJ427484	L13700
<i>Cba. tepidum</i>	-	-	AM690351	1287	industrially polluted fresh water pond, Hyderabad, India (17.39 N 78.49 E)	-	-	-	-
<i>Cba. thiosulfatophilum</i>	1430	-	AJ290825	1244	Federsee, Germany (48.08 N 9.63 E)	-	-	-	AJ391144
<i>Cba. thiosulfatophilum</i>	6230	DSM 249 <sup>T</sup>	Y08102	1419	Tassajara Hot Spring, California, USA (36.14 N 121.32 W)	AM049277*	AM050073*	AM050370*	AJ391143
<i>Cba. chlorovibrioides</i>	UdG 6026	-	Y10649	1466	Massona Lagoon (Parc Natural dels Aiguamolls de l'Empordà),	AM049279*	AM050066*	AM050352*	AJ306188/ AJ306182

Supplementary material

					Spain (42.21 N 3.10 E)				
<i>Cba. sp.</i>	UdG 6043	-	Y10648	1465	Everglades, Florida, USA (25.22 N 80.57 W)	AM049280*	AM050067*	AM050351*	AJ306193/ AJ306187
<i>Cba. sp.</i>	BS3	-	AY394784	717	Pichavaram Mangroves, Tamil Nadu, India (11.43 N 79.77 E)	-	-	-	-
<i>Cba. sp.</i>	JAGS72	-	FN662693	1348	Pichavaram Mangroves, Tamil Nadu, India (11.43 N 79.77 E)	-	-	-	-
<i>Cba. sp.</i>	C1	-	AM050125*	1440	Lake Waldsee, Germany (47.92 N 9.75 E)	AM049287*	AM050061*	AM050357*	*
<i>Cba. sp.</i>	JAGS57	-	FN662679	803	Rambha, Chilika Lagoon, Orissa, India (19.72 N 85.33 E)	-	-	-	-
<i>"Chl. bathyomarinum"</i>	GSB1	-	AY627756	1379	Juan de Fuca Ridge, Ty vent, Pacific Ocean (49.63 N 130.00 W)	AM049278*	AM050095*	AM050371*	AY627684
<i>Chl. chlorochromatii</i>	CaD3 <sup>T</sup>	-	AJ578461/Ca g_r1	1387/ 1509	Lake Dagow, Germany (53.15 N 13.05 E)	AM049317*	Cag_0387	Cag_1742	Cag_1329
<i>Chl. clathratiforme</i>	PG	-	Y08106	1355	Lake Grünenplan, Germany (51.94 N 9.79 E)	AM049281*	AM050093*	AM050367*	*
<i>Chl. clathratiforme</i>	BU-1	DSM 5477 <sup>T</sup>	Y08108/ Ppha_R0023	1308/ 1495	Lake Buchensee, Germany (50.48 N 10.15 E)	AM049282*	Ppha_2380	Ppha_0708	AJ290822/ Ppha_105 9

Supplementary material

<i>Chl. clathratiforme</i>	4DE	-	AM086645	1382	chemocline, Lake Cadagno, Switzerland (46.55 N 8.71 E)	-	-	-	-
<i>Chl. ferrooxidans</i>	KoFox	DSM 13031 <sup>T</sup>	Y18253/AAS E01000013	1495/ 1495	Ditch, Konstanz, Germany (47.68 N 9.17 E)	AM049283*	-	-	-
<i>Chl. gokarna</i>	4GSB <sup>T</sup>	-	AJ888464	1287	salt pan, Gokarna, India (14.55 N 74.31 E)	-	-	-	-
<i>Chl. limicola</i>	6330	DSM 245 <sup>T</sup>	Y10113/CP001097	1359/ 1494	Gilroy Hot Spring, California, USA (37.01 N 121.57 W)	AM049284*	Clim_0407	Clim_0630	AJ391153/ Clim_170 4
<i>Chl. limicola</i>	C2	-	ident. C1	1416	Lake Waldsee, Germany (47.92 N 9.75 E)	AM049286*	AM050063*	AM050356*	-
<i>Chl. limicola</i>	D1	-	ident. C1	1406	Lake Waldsee, Germany (47.92 N 9.75 E)	AM049289*	AM050064*	-	-
<i>Chl. limicola</i>	D2	-	ident. C1	-	Lake Waldsee, Germany (47.92 N 9.75 E)	AM049288*	AM050065*	-	-
<i>Chl. limicola</i>	E2P2	-	AM050126*	1432	Lake Sisó, Spain (42.13 N 2.76 E)	AM049290*	AM050082*	AM050359*	*
<i>Chl. limicola</i>	E3P1	-	ident. E2P2	-	Lake Sisó, Spain (42.13 N 2.76 E)	ident. E2P2*	AM050085*	AM050364*	*
<i>Chl. limicola</i>	E2P1	-	ident. E2P2	-	Lake Sisó, Spain (42.13 N 2.76 E)	ident. E2P2*	AM050080*	-	-
<i>Chl. limicola</i>	9330	-	AJ290827	1242	Clarke Reservation, New York, USA (43.00 N 76.08 W)	-	-	-	AJ391158
<i>Chl. limicola</i>	M1	-	AB054671	1465	-	-	-	-	-
<i>Chl. limicola</i>	UdG 6002	DSM 246	Y10640	1464	Pond, Meyershausen, Germany (51.61 N 9.99 E)	AM049285*, AM049291*	AM050072*	AM050355*	AJ391142

# Supplementary material

<i>Chl. limicola</i>	5230	DSM 1855	AJ290832	1240	Plußsee, Plön, Germany (54.18 N 10.44 E)	-	-	-	AJ391162
<i>Chl. limicola</i>	CNg	-	AM690805	1291	beach, marine tidal water, Narsapur, India (16.43 N 81.70 E)	-	-	-	-
<i>Chl. limicola</i>	Vz	-	AM690780	1286	sewage water near sea shore, Visakhapatnam, India (17.69 N 83.22 E)	-	-	-	-
<i>Chl. luteolum</i>	2530	DSM 273 <sup>T</sup>	Y08107/Plut_R0052	1403/ 1364	Lake Polden, Norway (60.71 N 5.03 E)	AM049296*	AM050090*/Plut_0387	AJ427490/Plut_1605	AJ391152/ Plut_1500
<i>Chl. luteolum</i>	E1P1	-	AM050131*	1419	Lake Sisó, Spain (42.13 N 2.76 E)	AM049311*	AM050074*	-	-
<i>Chl. luteolum</i>	E2P4	-	ident. E1P1	1377	Lake Sisó, Spain (42.13 N 2.76 E)	AM049312*	AM050075*	-	-
<i>Chl. luteolum</i>	E3P2	-	ident. E1P1	1446	Lake Sisó, Spain (42.13 N 2.76 E)	-	-	-	-
<i>Chl. phaeobacteroides</i>	2430	DSM 266 <sup>T</sup>	Y08104/Cpha266_R0053	1400/ 1494	Lake Blankvann, Norway (60.01 N 10.39 E)	AM049297*	Cpha266_2028	Cpha266_0583	AJ391148/ Cpha266_1888
<i>Chl. phaeobacteroides</i>	G-gCphb03	-	EF654662	1361	Lake Gek-Gel, Azerbaijan (40.41 N 46.33 E)	-	-	-	-
<i>Chl. phaeobacteroides</i>	1VIID7	-	AM086647	1381	chemocline, Lake Cadagno, Switzerland (46.55 N 8.71 E)	-	-	-	-
<i>Chl. phaeobacteroides</i>	CL1401	-	AM050127*	1350	Lake Sisó, Spain (42.13 N 2.76 E)	AM049298*	AM050077*	AM050362*	AJ427488
<i>Chl. phaeobacteroides</i>	E1P2	-	ident. CL1401	1438	Lake Sisó, Spain (42.13 N 2.76 E)	AM049299*	AM050079*	AM050360*	*
<i>Chl. phaeobacteroides</i>	E2P3	-	ident.	1411	Lake Sisó, Spain (42.13 N 2.76 E)	AM049300*	AM050081*	AM050361*	*

CL1401									
<i>Chl. phaeobacteroides</i>	E3P3	-	ident. CL1401	1447	Lake Sisó, Spain (42.13 N 2.76 E)	AM049301*, AM049302*	AM050084*	AM050363*	*
<i>Chl. phaeobacteroides</i>	Glu	-	AM050128*	1446	Lake Glaubitz, Germany (51.32 N 13.38 E)	AM049303*	AM050078*	AM050358*	*
<i>Chl. phaeobacteroides</i>	Dagow III	-	AM050129*	1356	Lake Dagow, Germany (53.15 N 13.05 E)	AM049304*	AM050086*	AM050365*	*
<i>Chl. phaeobacteroides</i>	BS-1	-	AJ972456/ Cphamn1_R0 003	1388/ 1495	Black Sea, Bulgaria (42.49 N 30.40 E)	AM039431*	Cphamn1_203 5	Cphamn1_06 02	Cphamn1_ 0841
<i>Chl. phaeovibrioides</i>	-	DSM 270	AJ290834	1240	Lake Polden, Norway (60.71 N 5.03 E)	-	-	-	AJ391164
<i>Chl. phaeovibrioides</i>	2631	DSM 269 <sup>T</sup>	Y08105	1466	Lake Landvikvannet, Norway (58.32 N 8.51 E) <sup>c</sup>	AM049308*	AM050091*	AJ427489	AJ391150
<i>Chl. phaeovibrioides</i>	2630	DSM 261	AJ290828	1241	Lake Landvikvannet, Norway (58.32 N 8.51 E) <sup>c</sup>	-	-	-	AJ391146
<i>Chl. phaeovibrioides</i>	1930	DSM 265	AJ290829/Cv ib_R0003	1241/ 1494	Sehestedt, Germany (54.37 N 9.82 E)	NC_009337, 193280-193777	Cvib_0443	Cvib_1395	AJ391160
<i>Chl. phaeovibrioides</i>	sy9	-	EU770420	1456	Qingdao Ocean, China (36.07 N 120.38 E)	-	-	-	-
<i>Chl. phaeovibrioides</i>	Mog4	-	EF149015	1414	Lake Mogilnoe, Kildin Island, Barents Sea (69.35 N 34.15 E)	-	-	-	-
<i>Chl. sp.</i>	6GSB	-	AJ888465	1371	salt pan, Sanikatta, India (14.51 N 74.36 E)	-	-	-	-

Supplementary material

<i>Chl. sp.</i>	ShCI03	-	EF153291	1364	Lake Shira, Khakassia (54.51 N 90.20 E)	-	-	-	-
<i>Chl. sp.</i>	200T-80	-	AB210277	969	activated sludge, South Korea (35.90 N 127.76 E)	-	-	-	-
<i>Chl. vibrioforme</i>	EP2403	-	AM050130*	1399	Prévost Lagoon, France (43.51 N 3.90 E)	AM049309*	-	-	-
<i>Chp. thalassium</i>	-	ATCC 35110 <sup>T</sup>	AF170103/Ct ha_R001	1403/ 1480	Woods Hole, Massachusetts, USA (41.53 N 70.67 W)	AM049310*	AM050088*/ Ctha_2619	Ctha_2034	Ctha_0710
<i>Chp.sp.</i>	16/3–127	-	DQ383304	1375	photobioreactor	-	-	-	-
<i>Chp. sp</i>	i9–5	-	DQ383317	1374	photobioreactor	-	-	-	-
<i>Ptc. aestuarii</i>	CE2401	-	AM050132*	1398	Certes lagoon, France (44.41 N 1.08 W)	AM049313*	AM050094*	AJ427493	-
<i>Ptc. aestuarii</i>	CHP3401	-	AJ291826	1458	Salada de Chiprana, Zaragoza, Aragon, Spain (41.23 N 0.18 W)	-	-	-	-
<i>Ptc. aestuarii</i>	SK-413	DSM 271 <sup>T</sup>	Y07837/ CP001108	1345/ 1495	Lagoon	AM049314*	Paes_1835	Paes_0545	AJ391151/ Paes_1595
<i>Ptc. aestuarii</i>	GG	-	AM690792	1288	marine tidal water, Visakhapatnam, India (17.69 N 83.22 E)	-	-	-	-
<i>Ptc. aestuarii</i>	FTGSB	-	AM690793	1289	marine tidal water, Visakhapatnam, India (17.69 N 83.22 E)	-	-	-	-
<i>Ptc. aestuarii</i>	P1	-	AM690797	1289	marine aquaculture pond water,	-	-	-	-

					Bheemli, India (17.98 N 83.22 E)				
<i>Ptc. aestuarii</i>	G48	-	AM690800	1292	beach, soil from saltpan, Goa, India (15.31 N 73.45 E)	-	-	-	-
<i>Ptc. aestuarii</i>	B7	-	AM690806	1296	marine aquaculture pond water, Bheemli, India (17.98 N 83.22 E)	-	-	-	-
<i>Ptc. aestuarii</i>	G25	-	AM690799	1283	soil from saltpan, Goa, India (15.20 N 73.54 E)	-	-	-	-
<i>Ptc. aestuarii</i>	sy10	-	EU770421	1457	marine water with sediment at 10m depth, Qingdao Ocean, China (36.07 N 120.38 E)	-	-	-	-
<i>Ptc. aestuarii</i>	UdG7004Chp	-	EF064309	1269	Salada de Chiprana, Zaragoza, Aragon, Spain (41.23 N 0.18 W)	-	-	-	-
<i>Ptc. indica</i>	4Kg	-	AM690781	1267	tidal water, Kakinada, India (16.95 N 82.24 E)	-	-	-	-
<i>Ptc. indica</i>	J29	-	AM690790	762	Kakinada, India (16.95 N 82.24 E)	-	-	-	-
<i>Ptc. indica</i>	6Pg	-	AM690783	1273	marine aquaculture pond water, Bheemli, India (17.98 N 83.22 E)	-	-	-	-
<i>Ptc. indicum</i>	SKGSB <sup>T</sup>		AJ887996	1393	Kakinada, India (16.95 N 82.24 E)	-	-	-	-
<i>Ptc. phaeum</i>	CIB2401	-	AJ291828	1453	Mallorca, Balears Island, Spain (39.69 N 3.02 E)	-	-	-	*
<i>Ptc. vibrioformis</i>	-	DSM 1678	AJ290833	1242	Seaside, Japan	-	-	-	AJ391163
<i>Ptc. vibrioformis</i>	JAGS42	-	FM177144	677	Orissa, Chilka lagoon, India (19.72 N 85.33 E)	-	-	-	-



# Supplementary material

<i>Ptc. vibrioformis</i>	ATCC 6030	DSM 260 <sup>T</sup>	M62791	1507	Moss Landing, California, USA (36.80 N 121.78 W)	AM049316*	AM050089*	AM050368*	AJ391145
<i>Ptc. vibrioformis</i>	CHP3402	-	AJ291827	1464	Salada de Chiprana, Zaragoza, Aragon, Spain (41.23 N 0.18 W)	-	-	-	AJ306185
<i>Ptc. vibrioformis</i>	3M	-	AM690791	1296	sediment from saltpan Machilipatnam, India (16.19 N 81.12 E)	-	-	-	-
<i>Ptc. vibrioformis</i>	UdG7006Lms	-	EF064312	1270	Massona Lagoon (Parc Natural dels Aiguamolls de l'Empordà), Spain (42.21 N 3.10 E)	-	-	-	-
<i>Ptc. vibrioformis</i>	UdG7009Lms	-	EF064313	1373	Massona Lagoon (Parc Natural dels Aiguamolls de l'Empordà), Spain (42.21 N 3.10 E)	-	-	-	-
<i>Ptc. vibrioformis</i>	UdG7008Cib	-	EF064315	1278	Cibollar Lagoon, Mallorca, Spain (39.70 N 3.01 E)	-	-	-	-
<i>Ptc. vibrioformis</i>	UdG7010Lms	-	EF064314	1376	Massona Lagoon (Parc Natural dels Aiguamolls de l'Empordà), Spain (42.21 N 3.10 E)	-	-	-	-
<i>Ptc. vibrioformis</i>	J18	-	AM690784	1281	sediment from saltpan, Kakinada, India (16.95 N 82.24 E)	-	-	-	-
<i>Ptc. sp.</i>	GSB2	-	HQ237491	502	Khakassia, salt Lake Shunet, Russia (54.42 N 90.22 E)	-	-	-	-
<i>Ptc. sp.</i>	JAGS24	-	FN543481	733	Kakinada, India (16.95 N 82.24 E)	-	-	-	-
<i>Ptc. sp.</i>	JAGS71	-	FN662683	779	Pichavaram Mangroves, Tamil	-	-	-	-

# Supplementary material

					Nadu, India (11.43 N 79.77 E)				
<i>Ptc. sp.</i>	JAGS73	-	FN662686	692	Tamil Nadu, Muthupettai Lagoon, India (10.33 N 79.55 E)	-	-	-	-
<i>Ptc. sp.</i>	JAGS113	-	FN984729	450	Rambha, Chilika Lagoon, Orissa, India (19.72 N 85.33 E)	-	-	-	-
<i>Ptc. sp.</i>	JAGS115	-	FN984731	410	Sholapatta, Orissa, India (20.95 N 85.10 E)	-	-	-	-
<i>Ptc. sp.</i>	5H2	-	AB301726	1397	microbial mat at a shallow submarine hot spring, Okinawa, off Taketomi island, Japan (24.19 N 124.05 E)	-	-	-	-
<i>Ptc. sp.</i>	JAGS111	-	FN773172	627	Tamil Nadu, Muthupettai Lagoon, India (10.33 N 79.55 E)	-	-	-	-
<i>Ptc. sp.</i>	2K	-	AJ290835	1240	-	AM049315*	-	AJ427492	AJ306196/ AJ290823
<i>Ptc. sp.</i>	ShNPel02	-	EF149016	1401	Khakassia, salt Lake Shunet, Russia (54.42 N 90.22 E)	-	-	-	-
" <i>Chlorobiaceae bacterium</i> "	LA53	-	AF513460	1460	Hawaii, USA (19.90 N 155.58 W)	-	-	-	-

*Cba. Chlorobaculum, Chl. Chlorobium, Pld. Pelodictyon, Ptc. Prosthecochloris*, n.i. not included, ident. identical, \*obtained during this study

Name of enrichment culture or env. sequence	Sequence name	Phylotype <sup>b</sup>	EMBL Accession No.	Length of sequence	Sampling site
<b>Phototrophic consortia<sup>a</sup></b>					
" <i>Chlorochromatium aggregatum</i> "	8a-2	N	AJ578422	482	Wintergreen Lake, Michigan, USA (42.340 N 85.38 W)
" <i>Chlorochromatium aggregatum</i> "	9b-2	D	AJ578405	471	Lake Sheffer, Michigan, USA (42.48 N 85.87 W)
" <i>Chlorochromatium aggregatum</i> "	CaC	P	AJ578433	487	Lake Coromina, Spain (42.13 N 2.56 E)
" <i>Chlorochromatium aggregatum</i> "	CaEl	J	AJ272091	448	Lake Echo, Washington, USA (47.04 N 121.42 W)
" <i>Chlorochromatium aggregatum</i> "	CaH	C	AJ578404	480	Lake Haussee, Germany (53.25 N 13.54 E)
" <i>Chlorochromatium aggregatum</i> "	CaS	P	AJ578432	487	Lake Sisó, Spain (42.13 N 2.76 E)
" <i>Chlorochromatium aggregatum</i> "	CaSp	L	AJ578418	491	Woods Hole, Massachusetts, USA (41.53 N 70.67 W)
" <i>Chlorochromatium glebulum</i> "	CGEIGSB5	F	AJ272092	448	Lake Echo, Washington, USA (47.04 N 121.42 W)
" <i>Chlorochromatium lunatum</i> "	12-2	D	AJ578408	480	Lake Cassidy, Michigan, USA (42.35 N 84.08 W)
" <i>Chlorochromatium magnum</i> "	10a	H	AJ578416	470	Lake Leach, Michigan, USA (42.69 N 85.29 W)
" <i>Chlorochromatium magnum</i> "	5-2	D	AJ578401	476	Mud Lake, Michigan, USA (46.19 N 84.11 W)
" <i>Chlorochromatium magnum</i> "	CmD	E	AJ272094	487	Lake Dagow, Germany (53.15 N 13.05 E)
" <i>Chlorochromatium magnum</i> "	CmElGSB2	H	AJ272093	448	Lake Echo, Washington, USA (47.04 N 121.42 W)
" <i>Pelochromatium roseum</i> "	18a	O	AJ578431	478	Lake Sheffer, Michigan, USA (42.48 N 85.87 W)
" <i>Pelochromatium roseum</i> "	18c-2	B	AJ578402	487	Lake Sheffer, Michigan, USA (42.48 N 85.87 W)

## Supplementary material

---

" <i>Pelochromatium roseum</i> "	25	K	AJ578417	480	Lake Schleinsee, Germany (47.61 N 9.63 E)
" <i>Pelochromatium roseum</i> "	PrDBd8	N	AY247958	470	Lake Haussee, Germany (53.09 N 13.03 E)
" <i>Pelochromatium roseum</i> "	PrEl	B	AJ578403	485	Lake Echo, Washington, USA (47.04 N 121.42 W)
" <i>Pelochromatium roseum</i> "	PrH	G	AJ578410	488	Lake Haussee, Germany (53.25 N 13.54 E)
" <i>Pelochromatium selenoides</i> "	PsH	D	AJ578409	488	Lake Haussee, Germany (53.25 N 13.54 E)

## Environmental clones and DGGE bands

5z-5	FJ797392	1368	gypsum karst lakes, Lithuania (56.06 N 24.40 E)
gsbspa11	AM050122*	451	Lake Kirkilai, Lithuania (56.24 N 24.69 E)
gsbspa12	AM050123*	458	Lake Kirkilai, Lithuania (56.24 N 24.69 E)
366	AJ831746	1381	Lake Cadagno, Switzerland (46.55 N 8.71 E)
810	AJ831751	1382	Lake Cadagno, Switzerland (46.55 N 8.71 E)
827	AJ831745	1382	Lake Cadagno, Switzerland (46.55 N 8.71 E)
849	AJ831742	1382	Lake Cadagno, Switzerland (46.55 N 8.71 E)
HH286	FJ502275	1471	Lake Cadagno, Switzerland (46.55 N 8.71 E)
LC15_L00C03	FJ547020	820	Lake Cadagno, Switzerland (46.55 N 8.71 E)
AL_BAC10	EU556482	446	Vestfold Hills, Ace Lake, Antarctica (68.28 S 78.19 E)
AL_BAC5	EU556478	485	Vestfold Hills, Ace Lake, Antarctica (68.28 S 78.19 E)
-	AY303358	426	Vestfold Hills, Ace Lake, Antarctica (68.28 S 78.19 E)
sediment_E5	EU910840	643	Lake Kinneret, Israel (32.49 N 35.35 W)
S_10	EU910848	753	Lake Kinneret, Israel (32.49 N 35.35 W)
L_8	EU910847	696	Lake Kinneret, Israel (32.49 N 35.35 W)

sediment_G7	EU910835	668	Lake Kinneret, Israel (32.49 N 35.35 W)
S_9	EU910851	758	Lake Kinneret, Israel (32.49 N 35.35 W)
L_5	EU910846	753	Lake Kinneret, Israel (32.49 N 35.35 W)
c5LKS19	AM086114	1423	Lake Kinneret, Israel (32.49 N 35.35 W)
s1ua1	DQ416209	740	
CAR8MG62	FJ902378	1187	coral <i>Oculina patagonica</i> , Sedot Yam, Israel (32.49 N 34.89 E) Caracol, Tamaulipas, Mexico (24.26 N 98.84 W)
CAR8MG90	FJ902402	1062	Caracol, Tamaulipas, Mexico (24.26 N 98.84 W)
CAR8MG68	FJ902384	808	Caracol, Tamaulipas, Mexico (24.26 N 98.84 W)
CAR8MG88	FJ902400	1123	Caracol, Tamaulipas, Mexico (24.26 N 98.84 W)
CAR8MG20	FJ902342	580	Caracol, Tamaulipas, Mexico (24.26 N 98.84 W)
CAR8MG94	FJ902406	832	Caracol, Tamaulipas, Mexico (24.26 N 98.84 W)
CAR8MG62	FJ902378	1187	Caracol, Tamaulipas, Mexico (24.26 N 98.84 W)
CAR8MG88	FJ902400	1123	Caracol, Tamaulipas, Mexico (24.26 N 98.84 W)
Zac7146	FJ485117	1291	El Zacaton, Mexico (22.80 N 102.04 W)
Zac4FP26	FJ484103	722	water column sample in El Zacaton, Mexico (22.80 N 102.04 W)
CC1_16S_7	EU662401	1409	microbial mat, Cesspool Cave, Virginia, USA (37.43 N 78.66 W)
CC1_16S_36	EU662387	1464	microbial mat, Cesspool Cave, Virginia, USA (37.43 N 78.66 W)
CC1B_16S_41	EU662315	1472	microbial mat, Cesspool Cave, Virginia, USA (37.43 N 78.66 W)
FGL12_B84,	FJ437858	1461	Green Lake, New York, USA (43.05 N 75.97 W)

---

FGL3_B75	FJ437773	1460	Green Lake, New York, USA (43.05 N 75.97 W)
K-100	AJ428422	1243	Sippewissett saltmarsh, Massachusetts, USA (41.58 N 70.61 W)
K-108	AJ428425	1241	Sippewissett saltmarsh, Massachusetts, USA (41.58 N 70.61 W)
K-137	AJ428426	1240	Sippewissett saltmarsh, Massachusetts, USA (41.58 N 70.61 W)
K-153	AJ428427	1240	Sippewissett saltmarsh, Massachusetts, USA (41.58 N 70.61 W)
K-701	AJ428430	1259	Sippewissett saltmarsh, Massachusetts, USA (41.58 N 70.61 W)
K-721	AJ428432	1259	Sippewissett saltmarsh, Massachusetts, USA (41.58 N 70.61 W)
K-902	AJ428433	1261	Salt lake mud, Death Valley, California, USA (36.38 N 116.90 W)
MV4-01-ENVL	AY394778	753	thermal spring, Yellowstone National Park, USA (44.46 N 110.64 W)
MV2-00-env2	AY394835	432	thermal spring, Yellowstone National Park, USA (44.46 N 110.64 W)
BS1-01-env9	AY394802	512	thermal spring, Yellowstone National Park, USA (44.46 N 110.64 W)
MV1-00-env6	AY394752	512	thermal spring, Yellowstone National Park, USA (44.46 N 110.64 W)
MV6-02-env4	AY394842	518	thermal spring, Yellowstone National Park, USA (44.46 N 110.64 W)
BS1-00-env3	AY394805	512	thermal spring, Yellowstone National Park, USA (44.46 N 110.64 W)
A4S16P_ML_093	FJ352497	849	New Orleans, USA (29.95 N 90.07 W)
A4S16P_ML_215	FJ352406	849	New Orleans, USA (29.95 N 90.07 W)
A4S16S_ML_297	FJ352694	849	New Orleans, USA (29.95 N 90.07 W)
A4S16P_ML_189	FJ352498	790	New Orleans, USA (29.95 N 90.07 W)
G912P35RE12	EU171708	842	Boulder, Colorado, USA (40.01 N 105.27 W)

---

L69_Loihi	AY371184	754	Liohi submarine volcano, Hawaii, USA (19.90 N 155.58 W)
K2-4-2	AY344404	1461	Lake Kauhakō, Hawaii, USA (19.90 N 155.58 W)
L125-11	AY344407	860	Lake Kauhako, Hawaii, USA (19.90 N 155.58 W)
Zplanct25	EF602486	706	Zodletone Spring, South Carolina, USA (32.88 N 80.66 W)
ZB24	AY327197	783	Zodletone Spring, South Carolina, USA (32.88 N 80.66 W)
ZB22	AY327195	788	Zodletone Spring, South Carolina, USA (32.88 N 80.66 W)
SAW1_B49	FJ716275	1462	Sawmill Sink, Bahamas, USA (25.03 N 77.40 W)
SAW1_B66	FJ716339	1466	Sawmill Sink, Bahamas, USA (25.03 N 77.40 W)
SAW1_B68	FJ716337	1461	Sawmill Sink, Bahamas, USA (25.03 N 77.40 W)
SAW1_B70	FJ716336	1462	Sawmill Sink, Bahamas, USA (25.03 N 77.40 W)
SAW1_B15	FJ716306	1463	Sawmill Sink, Bahamas, USA (25.03 N 77.40 W)
SAW1_B21	FJ716300	1462	Sawmill Sink, Bahamas, USA (25.03 N 77.40 W)
Dago99a	AM050117*	477	Lake Dagow, Germany (53.15 N 13.05 E)
GSB5	AJ006182	443	Lake Dagow, Germany (53.15 N 13.05 E)
GSB7	AJ006184	442	Lake Dagow, Germany (53.15 N 13.05 E)
Dagow99a	AY247960	497	Lake Dagow, Germany (53.15 N 13.05 E)
FSW-St05-2	GU734037	574	Lake Große Fuchskuhle, Germany (53.25 N 13.08 E)
Fuku1-SW24	EF394801	433	Lake Große Fuchskuhle, Germany (53.25 N 13.08 E)
K-500	AJ428429	1259	Beach mud, Kiel bight, Baltic Sea, Germany (54.23 N 10.12 E)



---

ME-JP105	AJ784965	679	Mariager Fjord, Denmark (56.40 N 9.59 E)
ME-TG114	AJ784975	446	Mariager Fjord, Denmark (56.40 N 9.59 E)
ME-TG116	AJ784977	450	Mariager Fjord, Denmark (56.40 N 9.59 E)
MW-JP101	AJ748280	401	Mariager Fjord, Denmark (56.40 N 9.59 E)
MW-JP111	AJ748286	401	Mariager Fjord, Denmark (56.40 N 9.59 E)
SI_02	AJ580958	465	Lake Sisó, Spain (42.13 N 2.76 E)
SI_03	AJ580959	468	Lake Sisó, Spain (42.13 N 2.76 E)
SI_04	AJ580960	468	Lake Sisó, Spain (42.13 N 2.76 E)
SI_05	AJ580961	461	Lake Sisó, Spain (42.13 N 2.76 E)
SI_07	AJ580963	461	Lake Sisó, Spain (42.13 N 2.76 E)
SI_08	AJ580964	466	Lake Sisó, Spain (42.13 N 2.76 E)
Sisó1	AJ809887	473	Lake Sisó, Spain (42.13 N 2.76 E)
Sisó2	AJ809888	476	Lake Sisó, Spain (42.13 N 2.76 E)
1B5	AJ627984	650	Ebro Delta, Spain (40.65 N 0.75 E)
Coromina	AJ809885	473	Lake Coromina, Spain (42.13 N 2.56 E)
gsbspa1	AM050118*	462	Lake Coromina, Spain (42.13 N 2.56 E)
gsbspa2	AM050119*	459	Lake Negre, Spain (41.99 N 1.52 E)
TOBACB35	AM749878	516	El Tobar, Cuenca, Spain (40.55 N 2.06 W)
ARBAC34	AM749854	514	El Tobar, Cuenca, Spain (40.55 N 2.06 W)

---

Chp-5	DQ984167	452	Salada de Chiprana, Zaragoza, Aragon, Spain (41.23 N 0.18 W)
M-9	DQ984151	503	Massona Lagoon (Parc Natural dels Aiguamolls de l'Empordà), Spain (42.21 N 3.10 E)
M-11	DQ984153	507	Massona Lagoon (Parc Natural dels Aiguamolls de l'Empordà), Spain (42.21 N 3.10 E)
M-6	DQ984148	504	Massona Lagoon (Parc Natural dels Aiguamolls de l'Empordà), Spain (42.21 N 3.10 E)
M-4	DQ984146	504	Massona Lagoon (Parc Natural dels Aiguamolls de l'Empordà), Spain (42.21 N 3.10 E)
M-8	DQ984150	502	Massona Lagoon (Parc Natural dels Aiguamolls de l'Empordà), Spain (42.21 N 3.10 E)
M-18	DQ984160	542	Massona Lagoon (Parc Natural dels Aiguamolls de l'Empordà), Spain (42.21 N 3.10 E)
PI-5	DQ984172	408	Playa lagoon (Ebro Depression), Spain (37.94 N 0.78 W)
O-2	EF064307	516	sediment sample, Onyar River, Spain (41.88 N 2.74 E)
O-1	EF064306	535	sediment sample, Onyar River, Spain (41.88 N 2.74 E)
VIBAC6Unid	AJ240010	526	Lake Vilar, Spain (42.08 N 2.44 E)
Vilar	AJ809890	471	Lake Vilar, Spain (42.08 N 2.44 E)
WETLE-25R	FM992029	535	Léon, Spain (42.60 N 5.57 W)
Redon_L6-B4-T7	FN296755	923	Lake Redon water column, Spain (42.67 N 1.00 E)
Font de la Puda	AJ809886	473	Font de la Puda (Pudosa), Spain (42.11 N 2.75 E)

---

Thp_B_45	EF444751	1175	Thermopiles hot springs, Greece (38.80 N 22.54 E)
Nit5Au0613_180	EU570902	1472	Nitinat Lake, Canada (48.74 N 124.75 W)
Nit5A0622_517	FJ628264	1462	Nitinat Lake, Canada (48.74 N 124.75 W)
Nit5A0622_525	FJ628267	1464	Nitinat Lake, Canada (48.74 N 124.75 W)
Nit5A0622_505	FJ628258	1463	Nitinat Lake, Canada (48.74 N 124.75 W)
Nit5A0622_495	FJ628254	1463	Nitinat Lake, Canada (48.74 N 124.75 W)
Nit2A0626_75	EU266000	1461	Nitinat Lake, Canada (48.74 N 124.75 W)
Nit2A0626_83	EU266008	1461	Nitinat Lake, Canada (48.74 N 124.75 W)
Nit2A0626_239	FJ628193	1463	Nitinat Lake, Canada (48.74 N 124.75 W)
Nit2A0626_207	FJ628181	1475	Nitinat Lake, Canada (48.74 N 124.75 W)
Nit5A0622_515	FJ628263	1462	Nitinat Lake, Canada (48.74 N 124.75 W)
SLB-W-B13	EF614441	541	St. Lucia Bay, filtered seawater (13.76 N 60.97 W)
SLB-W-B18	EF614442	549	St. Lucia Bay, filtered seawater (13.76 N 60.97 W)
A133-21	AY323159	500	white band diseased <i>Acropora palmata</i> , Barbados (13.19 N 59.54 W)
MekkoH-29	AM949432	421	Mekkojarvi , Finland (61.37 N 25.15 E)
MekkoH-137	AM949396	437	Mekkojarvi , Finland (61.37 N 25.15 E)
MekkoM1-78	AM949229	452	Mekkojarvi , Finland (61.37 N 25.15 E)
LS DGGE band 3	AB469923	527	Fukui, Lake Suigetsu, Japan (36.06 N 136.21 E)
LS DGGE band 36	AB469945	574	Fukui, Lake Suigetsu, Japan (36.06 N 136.21 E)

---

---

N-L-1	AB154488	485	Lake Kaiike, Kamikoshiki Island, Japan (31.83N 129.87 E)
26s	AB154457	539	Lake Kaiike, Kamikoshiki Island, Japan (31.83N 129.87 E)
25s	AB154456	535	Lake Kaiike, Kamikoshiki Island, Japan (31.83N 129.87 E)
21ws	AB154452	539	Lake Kaiike, Kamikoshiki Island, Japan (31.83N 129.87 E)
23ws	AB154454	539	Lake Kaiike, Kamikoshiki Island, Japan (31.83N 129.87 E)
pItb-HW-48	AB294906	980	Okinawa, off Taketomi island, Japan (24.33 N 124.09 E)
LCK-32 16S	AJ296585	438	Petit Saut Lake, French Guiana (5.36 N 52.95 W)
LEchl-2	AJ296584	552	Petit Saut Lake, French Guiana (5.36 N 52.95 W)
S29	AF356020	711	White Island, New Zealand (37.52 S 177.18 E)
JAGS87	FN675246	439	Tamil Nadu, Thazrangadu, India (13.15 N 80.29 E)
JAGS88	FN675247	437	Tamil Nadu, Thazrangadu, India (13.15 N 80.29 E)
JAGS89	FN675248	437	Tamil Nadu, Thazrangadu, India (13.15 N 80.29 E)
JAGS84	FN675255	439	Tamil Nadu, Nagapattinam, India (10.77 N 79.83 E)
JAGS86	FN675257	441	Tamil Nadu, Nagapattinam, India (10.77 N 79.83 E)
JAGS85	FN675256	433	Tamil Nadu, Nagapattinam, India (10.77 N 79.83 E)
JAGS99	FN773174	430	Tamil Nadu, Kaipanikuppam, India (13.04 N 80.19 E)
JAGS110	FN773193	435	Tamil Nadu, Uppur, India (10.32 N 76.90 E)
JAGS107	FN773190	414	Tamil Nadu, Uppur, India (10.32 N 76.90 E)
JAGS61	FN662689	561	Satpada, Chilika Lagoon, Orissa, India (19.72 N 85.33 E)

---

---

JAGS74	FN675249	442	Satpada, Chilika Lagoon, Orissa, India (19.72 N 85.33 E)
JAGS91	FN773183	435	Satpada, Chilika Lagoon, Orissa, India (19.72 N 85.33 E)
JAGS90	FN773182	429	Satpada, Chilika Lagoon, Orissa, India (19.72 N 85.33 E)
JAGS77	FN675252	410	Satpada, Chilika Lagoon, Orissa, India (19.72 N 85.33 E)
JAGS96	FN773188	424	Satpada, Chilika Lagoon, Orissa, India (19.72 N 85.33 E)
JAGS97	FN773189	421	Satpada, Chilika Lagoon, Orissa, India (19.72 N 85.33 E)
JAGS76	FN675251	441	Satpada, Chilika Lagoon, Orissa, India (19.72 N 85.33 E)
JAGS79	FN675259	434	Satpada, Chilika Lagoon, Orissa, India (19.72 N 85.33 E)
JAGS93	FN773185	425	Satpada, Chilika Lagoon, Orissa, India (19.72 N 85.33 E)
JAGS81	FN675261	433	Satpada, Chilika Lagoon, Orissa, India (19.72 N 85.33 E)
JAGS80	FN675260	440	Satpada, Chilika Lagoon, Orissa, India (19.72 N 85.33 E)
JAGS106	FN773181	432	Tamil Nadu, Muthupettai Lagoon, India (10.33 N 79.55 E)
JAGS104	FN773179	433	Tamil Nadu, Muthupettai Lagoon, India (10.33 N 79.55 E)
JAGS103	FN773178	442	Tamil Nadu, Muthupettai Lagoon, India (10.33 N 79.55 E)
JAGS101	FN773176	438	Tamil Nadu, Muthupettai Lagoon, India (10.33 N 79.55 E)
JAGS98	FN773173	422	Tamil Nadu, Kaipanikuppam, India (13.04 N 80.18 E)
JAGS82	FN675253	432	West Bengal, Biraj Nagar, Sunderbans, India (21.95 N 88.86 E)
BS_sed_H	AM779897	494	Black Sea, Bulgaria (41.91 N 3.03 E)
BS_sed_D	AM779893	451	Black Sea, Bulgaria (41.91 N 3.03 E)

---

BS130	GU145513	1465	Black Sea (42.49 N 30.40 E)
BS061	GU145446	1463	Black Sea (42.49 N 30.40 E)
BS140	GU145523	1465	Black Sea (42.49 N 30.40 E)
R3ENDE8	DQ401523	921	wastewater, National university of Ireland (53.28 N 9.06 W)
R3ENDE14	DQ401529	884	wastewater, National university of Ireland (53.28 N 9.06 W)
Kis06_53	EU428901	812	water purification plant, Veendam, Netherlands (53.11 N 6.86 E)
B19	GU557111	600	wastewater treatment plant, Jiangmen city, China (22.58 N 113.08 E)
MTW4	EU275405	448	Lake Matano, Sulawesi Island, Indonesia (2.49 N 121.38 E)
LMT3	EU275404	415	Lake Matano, Sulawesi Island, Indonesia (2.49 N 121.38 E)
1HP1-A11	EU780307	800	<i>Turbinaria mesenterina</i> colony, Australia (27.96 N 153.41 E)
E03P2MbD	FJ156429	428	diseased coral colony Abrolhos, Bahia, Brazil (18.03 N 38.67 W)
CD4H2	AY038478	611	Corals, Curaçao, Netherlands Antilles (12.21 N 68.95 W)
CD5B9	AY038426	622	Corals, Curaçao, Netherlands Antilles (12.21 N 68.95 W)
S6D	AF298534	477	Mediterranean sapropel layers (41.64 N 8.17 E)
S8F	AF298536	477	Mediterranean sapropel layers (41.64 N 8.17 E)
S8G	AF298537	477	Mediterranean sapropel layers (41.64 N 8.17 E)
Z1C	AF298533	477	Mediterranean sapropel layers (41.64 N 8.17 E)
1_2-B9	FN824920	1262	groundwater treatment system Germany
CV-9	FN994991	1472	chemocline enrichment culture

C2	EU478457	452	contaminated sediment, Myllykoski, River Kymijoki, Finland (60.76 N 26.77 E)
D1T_073	EF444200	547	freshwater wetland soils
2.30	GQ183244	760	constructed treatment wetland mesocosms
dairy_MEC12	HM149049	730	dairy wastewater
CARB_ER1_22	AY239565	820	whey-based wastewater
D1-31	AF544880	450	Caribbean coral <i>Montastrea annularis</i>
D1-8	AF544872	450	Caribbean coral <i>Montastrea annularis</i>
AME E34	DQ191704	919	anaerobic wastewater bioreactor , National University of Ireland (53.28 N 9.06 W)?
AME E19	DQ191736	635	anaerobic wastewater bioreactor, National University of Ireland (53.28 N 9.06 W)?
16S6	DQ514570	441	bioreactor
16S4	DQ514568	440	bioreactor
16/3–165	DQ383307	1372	photobioreactor, University of Brisbane (27.49 S, 153.02 E)
16/3–107	DQ383298	1259	photobioreactor, University of Brisbane (27.49 S, 153.02 E)
16/3–116	DQ383301	1380	photobioreactor, University of Brisbane (27.49 S, 153.02 E)
16/3–163	DQ383305	1465	photobioreactor, University of Brisbane (27.49 S, 153.02 E)
i9–106	DQ383316	1385	photobioreactor, University of Brisbane (27.49 S, 153.02 E)
GSB6	AJ006183	442	-

## Supplementary material

---

---

CARB_ER1_22	AF107326	476	-
PSBAC-5	AJ698077	472	-
Achl-4	AJ698081	470	-
Achl-5	AJ698078	476	-
CIBAC-3	AJ239993	548	-

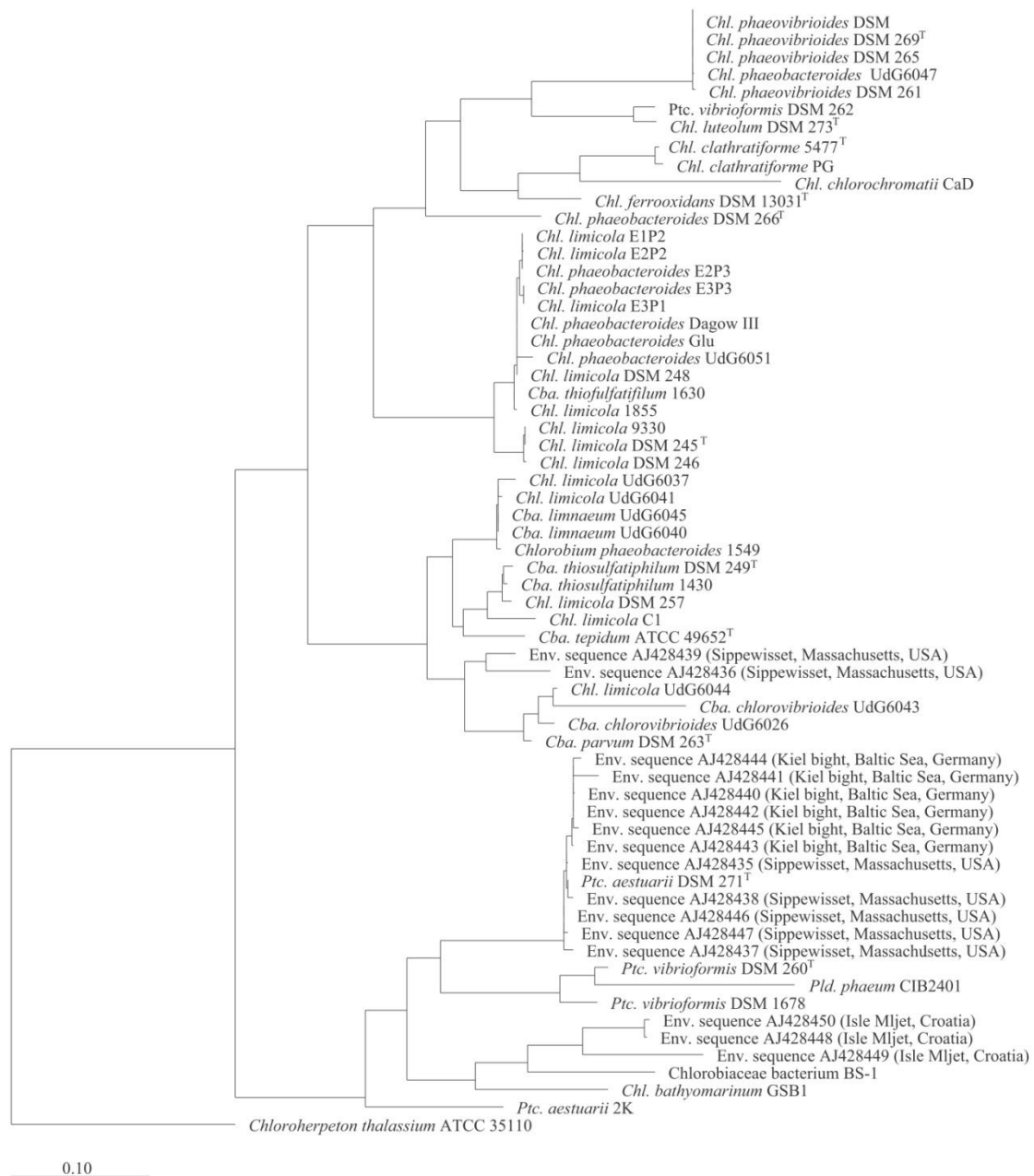
---

<sup>a</sup> picked manually from enrichment culture

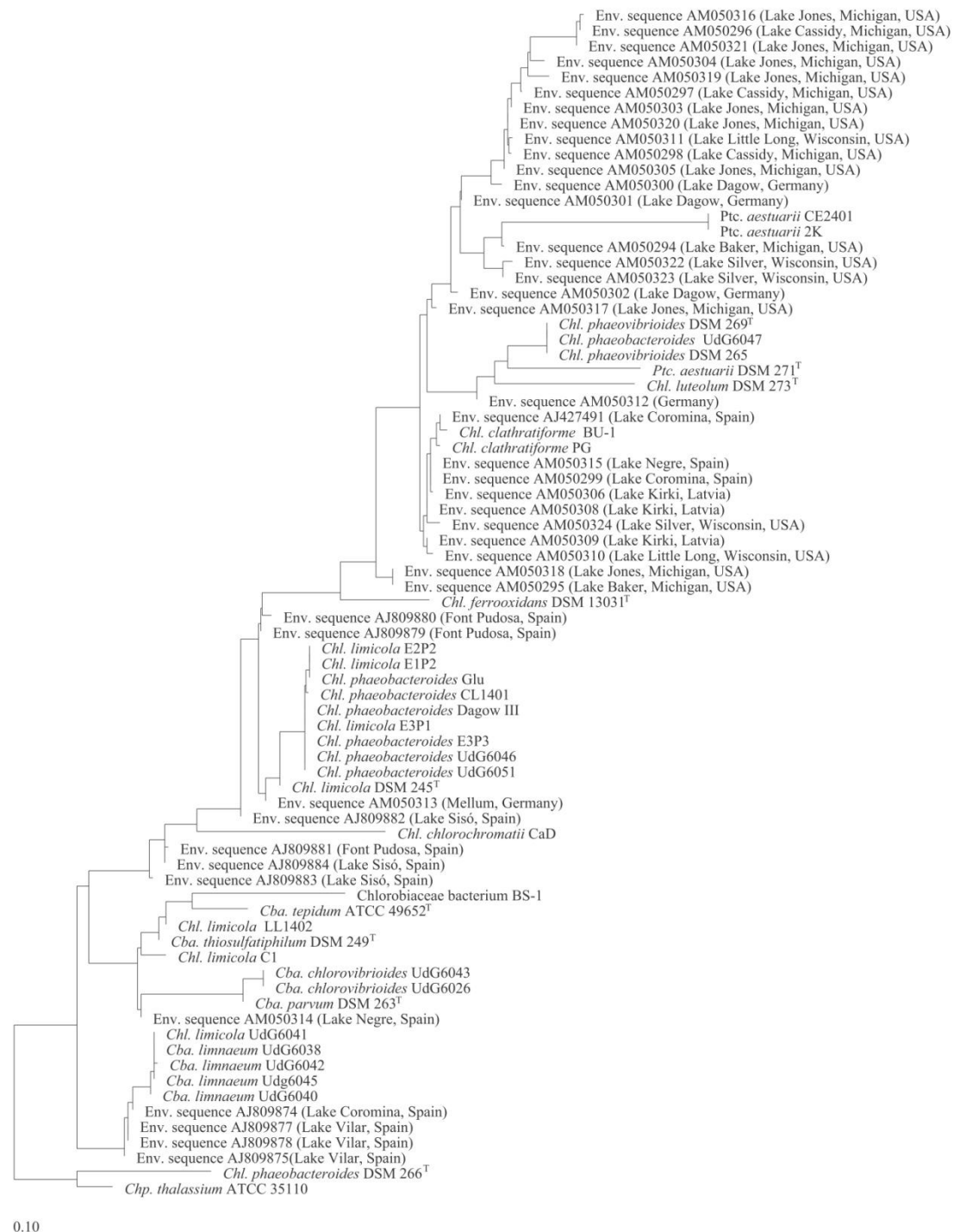
<sup>b</sup> Glaeser and Overmann, 2004

<sup>c</sup> This is opposed to the original entry “Lake Langvikvann” that does not exist. Lake Landvikvannet matches “Lake Langvikvann” in salt concentration depth of the anoxic zone and supposed location (southern Norway).

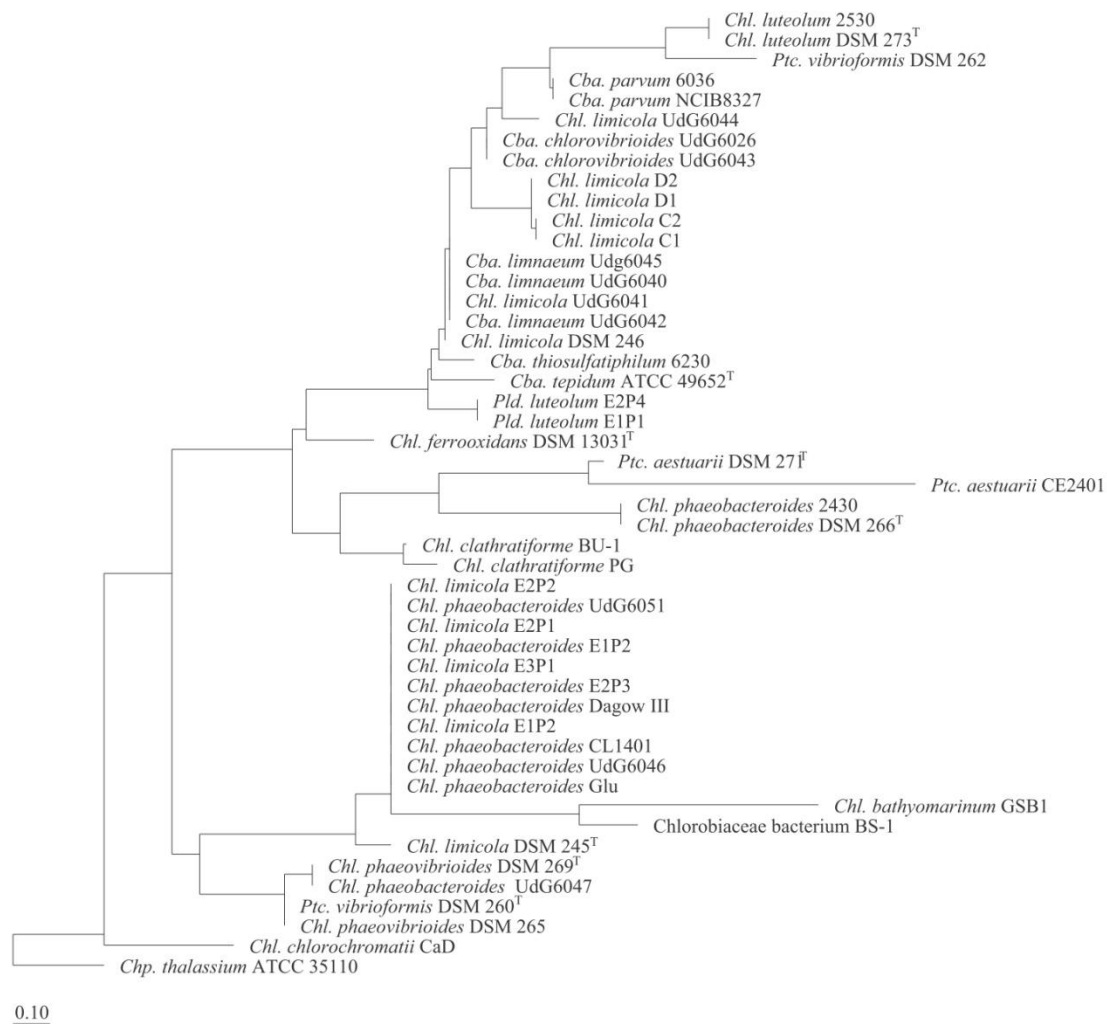




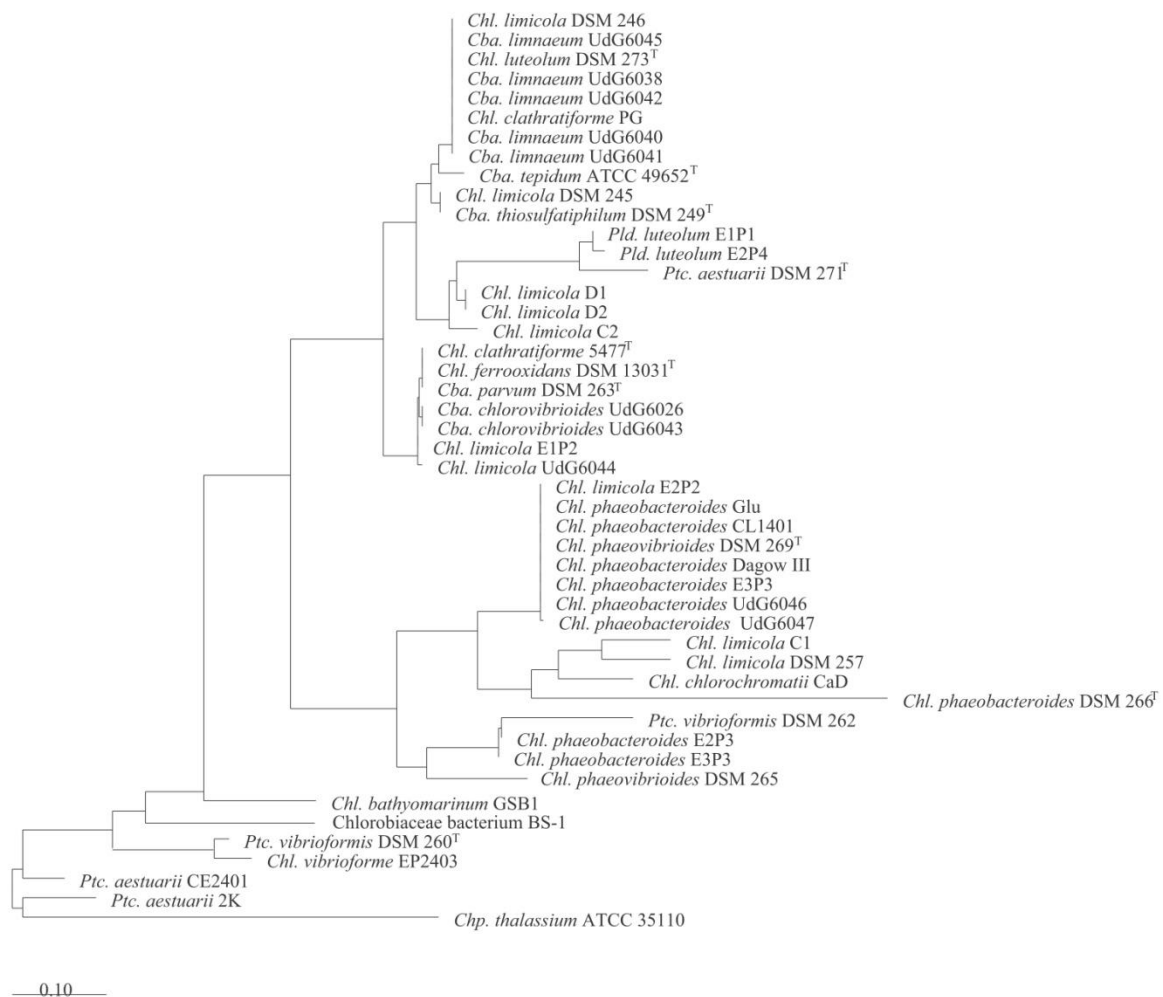
**Suppl. Fig. S1.** Phylogenetic tree of the available *fmoA* gene sequence of GSB. The tree was constructed in ARB, using the FAST\_DNAML algorithm. Bar denotes fixed substitutions per nucleotide.



**Suppl. Fig. S2.** Phylogenetic tree of the available *bchG* gene sequence of GSB. The tree was constructed in ARB, using the FAST\_DNAML algorithm. Bar denotes fixed substitutions per nucleotide.



**Suppl. Fig. S3.** Phylogenetic tree of the available *sigA* gene sequence of GSB. The tree was constructed in ARB, using the FAST\_DNAML algorithm. Bar denotes fixed substitutions per nucleotide.



**Suppl. Fig. S4.** Phylogenetic tree of the available 16S-23S rRNA intergenic spacer (ITS) region of GSB. The tree was constructed in ARB, using the FAST\_DNAML algorithm. Bar denotes fixed substitutions per nucleotide.